Development and Application of Computational Tools to Study Optical Rotation of Chiral Molecules in Isotropic Media

Tal Aharon

Submitted to the graduate degree program in Department of Chemistry and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Marco Caricato, Chairperson

Brian B. Laird

Committee members

Christopher G. Elles

Cindy L. Berrie

Hartwin Peelaers

Date defended: August 30, 2019

The Dissertation Committee for Tal Aharon certifies
that this is the approved version of the following dissertation :

Development and Application of Computational Tools to Study Optical Rotation of Chiral Molecules in Isotropic Media

Marco Caricato, Chairperson

Date approved: September 6, 2019

Abstract

The study of optically active molecules has been continuous for over two centuries, due to the role chiral molecules play as both the building blocks of life, and therefore their importance in developing pharmaceuticals. Theoretical calculations to study chiroptical properties, such as optical rotation (OR), electronic circular dichroism (ECD), vibrational circular dichroism (VCD), and Raman optical activity (ROA), have only been developed in the last two decades. Since then, two fundamental questions have dominated research of chiral molecules. The first question lies in successfully relating the magnitude and sign of chiral properties to a structure, as there is no chemically intuitive way to determine, without performing measurements or calculations, if the OR of a given chiral molecule will be small or large, positive or negative. The second relates to the effect of solvation, as we currently do not know the root causes of the changes induced in the OR by a change in the environment.

We investigate the structure-property relationship by examining helicenes, an interesting class of chiral molecules due to their intrinsic structural chirality (an unfunctualized helicene has no chiral centers) and their large OR. We used a method that decomposes the calculated OR into occupied-virtual MO pairs, and allows us to investigate what drives the large chiroptical response of helicenes. We examined the effect of the pitch, which is the distance between one full turn of a helix, and the length of helicenes. We then functionalized them with electron withdrawing and electron donating groups to investigate how this changed the OR. We find that helicene OR is driven by large magnetic dipoles and that we can characterize the type of contributions to OR into three subgroups by the general angle of the magnetic dipole with respect to the helical axis. We also extend our method to localized orbitals, which present the orbitals in the more chemically intuitive form of bonds and lone pairs.

Next, we investigated solvation effects in concert with the Vaccaro group at Yale University.

They measured gas and solution phase OR values for a series of rigid chiral molecules. We wanted to investigate how capable the polarizable continuum model (PCM) was at reproducing solvent effects and trends on OR. We performed calculations at a series of wavelengths and in a series of solvents, and compared our results to the calculated values. Gas phase calculations agreed very well with experimental values; however, PCM was unable to reproduce not only the solvent shifts, but also the trends in solvent shifts. We then calculated zero point vibrational corrections (ZPVCs) to the OR in an effort to improve our results, with mixed success.

We also worked to reduce the cost of computational calculations of OR, as cheaper calculations would allow us to investigate larger or more complicated systems, including better descriptions of solvent effects. Our efforts were two-fold. First, we developed a method to reduce the cost of the linear response equations by removing a portion of the orbitals prior to the solution of the coupled perturbed equations. We tested a variety of cutoff types and magnitudes with two basis sets and two functionals, and found that despite discarding large portions of the MOs, we maintained high accuracy and showed large speedups. Next, we optimized the exponents of a small basis set for OR calculations by taking the diffuse functions of large basis sets and appending them to a smaller one. We performed multiple tests with our new basis sets on a variety of molecules with OR ranging from $<$10 to $>$4,000 deg dm^{-1} (g/mL)$^{-1}$. The resulting basis sets were highly accurate, capable of producing values similar to the larger basis sets at a reduced cost.

Finally, we collaborated with the Avarvari group at the Université d'Angers, who study functionalized helicenes. They were interested in synthesizing helicenes that could act as circularly polarized light (CPL) emitters. They began by synthesizing for the first time helicenes with a strong electron withdrawing group (benzothiadiazole, BTD), which they found were not emissive. We performed calculations of electronic and chiroptical spectra to characterize the transitions, and were able to show that the fluorescence quenching was likely due to intersystem crossing in their helicenes. Their next series of helicenes were highly emissive, making them likely candidates for CPL emitters, and due to the agreement with our calculated spectra, we were again able to characterize the transitions.

Acknowledgements

I want to begin by thanking my advisor, Dr. Marco Caricato, for taking a chance on me from the very beginning. Since then, he has been incredibly patient and supportive, and I could not have gotten this far without his help and advice. He has gone out of his way to foster a positive atmosphere within the group, which has led us to support each other.

I would like to thank my group, both current members: Matt Barclay, Sijin Ren, Amy Jystad, Ty Balduf, Kaihua Zhang, Katherine Vander Laan, and Ryan Lampton, and past members: Dr. Alessandro Biancardi, Robert Adams, Hunter LeBlanc, Clara Stanton, Delaney Lynam, Analila Valenzuela, Jeremy Barnes, Joseph Harms, and Claudiu Caraiani. They have been a great resource when I needed to discuss research, and without them the work in this dissertation would not have been possible. They have also been a great support system, making them part of the foundation which has allowed me to succeed here at KU.

Next, I would like to thank my family. They have been there for me the entirety of this experience, encouraging me when I needed it but reminding me to take a break when I would forget. I feel very lucky to have them in my life.

Finally, I need to thank my friends. I have grown a lot in my time at KU, and I feel a large portion of that is due to them. I do not think I would have succeeded at my time here without them, and I am eternally grateful to have met such wonderful people. Especially important are those who I started with: Nate, Josh, Chris, and Matt. These 4 have supported me from the start, helped me grow and without them I would have never been able to get to know Zeke, Kelci, Amy, Kristen, David, Rob, and Liz as well as I have.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

A molecule is chiral if its mirror image is not superimposible on the original structure. The two enantiomers behave identically in achiral environments, and thus have the same boiling points, density, NMR spectra, etc., but interact differently when exposed to a chiral environment, e.g. biological systems. For instance, (R)-limonene is known for its lemon-like smell and is commonly used in household cleaning agents, whereas (L)-limonene has a turpentine-like smell, and is often used in industrial cleaning agents.[1] We can recognize the two odors because our nose is a chiral environment. In fact, all of life, including the human body, is built with only one enantiomer of the amino acids in the proteins and the sugar bases in the DNA, namely L-amino acids and D-sugars. This phenomenon is known as homochirality, and its origin is still not well understood.[2] As many pharmaceutically active molecules are also chiral, it is incredibly important to take proper precautions when testing and selling chiral drugs.[3] Because the efficacy of a chiral drug is directly related to the enantiomer, the assignment of absolute configuration (AC) is of paramount importance. In the past, AC was often assigned via ^{1}H NMR or X-ray crystallography, although these methods were time consuming and came with no guarantee of success.[4] This led to the popularity of spectroscopic techniques to identify chiral molecules, which measure a property that is directly tied to the structure of the molecule, for example optical rotation (OR), electronic circular dichroism (ECD), or most commonly vibrational circular dichroism (VCD). However, these techniques do not provide any indication as to which structure produces a certain response.

Because experimental spectroscopy cannot solely assign the AC, the development of accurate quantum chemical methods for simulating chiroptical properties has led to a synergistic approach where calculations are paired with experiments.[2] The first specific rotation calculation (usually

denoted as $[\alpha]_\omega$) was performed by Polavarapu in 1997 at the Hartree-Fock level of theory.[5] In the last two decades, optical rotation calculations were extended to modern density function theory (DFT) and coupled cluster (CC) theory.[6–20]

Despite many advances, a full understanding of the nature of the OR eludes researchers, as there remain two open ended topics in the study of chiral properties. The first is related to the intrinsic relationship between the optical activity of a molecule and the structure, as there remains no intuitive explanation as to why certain structures lead to large or small $[\alpha]_\omega$, or one sign in preference to the other. The second is related to the solution phase; there is a non-obvious effect the environment plays upon optical activity, as it is not straightforwardly related to the polarity of the solvent.[2]

Studies of the intrinsic structure-property relationship between chiral molecules and OR have seen work to decompose the equations used to calculate optical rotation[21–25]. Conformational and vibrational effects play a pivotal role on the magnitude and sign of $[\alpha]_\omega$, as chiral molecules are extremely sensitive to the geometry,[12,26–34] and that zero point vibrational corrections (ZPVCs)[35–44] can constitute a large part of the $[\alpha]_\omega$.

Solvent effects have seen just as much if not more interest, with the development of gas phase experimental measurements.[45] PCM, a simple solvation model that has seen effective use in reproducing solvent effects on other physical properties, has been tested and found to be unable to reproduce solvent specific rotation values.[27,34,46–53] Crawford *et al.* have performed calculations of solvent effects using chiral molecules embedded in a frozen density potential of a solvent system,[54] while Haghdani *et al.* used quantum mechanical micro-solvated clusters in an attempt to reproduce solvent effects. Molecular-dynamics simulations have been used to generate snapshots of amino acid and solvent systems, which were used to calculate $[\alpha]_\omega$ by using quantum chemical methods to treat the solute and treating the solvent as a series of point charges.[55,56] Beratan and co-workers performed similar calculations; however, the edge of their point charge system is surrounded by PCM.[57,58] Capelli and co-workers have developed a polarizable embedding scheme in which the solvent is treated by a series of both point and polarizable charges, which they call

the fluctuating charge method.[59,60] While these methods have produced correct results in some calculations, they are not consistently able to predict either the magnitude of solvent shifts from the gas phase or determine which solvent induces the largest change in the $[\alpha]_\omega$ when the molecule is solvated.

Optical rotation studies are made even more difficult due to the cost of the calculations, which comes from the need for extensive treatment of electron correlation effects, as the property is sensitive to the entire electronic spectrum of a molecule and not localized to the effect of small chromophores, and large basis sets, because optical activity is sensitive to the periphery of the wavefunction.[2,61] Combating this expense has therefore become important, as we seek to study new and larger systems. A basis set has been developed specifically for calculating optical rotation that can produce $[\alpha]_\omega$ values calculated with aug-cc-pVTZ at a cost closer to aug-cc-pVDZ.[62,63] By appending the diffuse functions from a large basis set to a much smaller basis set, Wiberg *et al.* showed that a small basis set with diffuse functions could be used to predict the $[\alpha]_\omega$ calculated with larger basis sets.[61] Additionally, local correlation methods have been investigated to alleviate some of the cost from expensive electronic structure methods.[64–66]

As a particularly interesting class of chiral molecules due to large $[\alpha]_\omega$ and an intrinsic structural chirality, helicenes have been a topic of continuous research for the last century.[67–69] They have a variety of applications, including use in LCD and OLED screens,[70–72] organic transistors[73] and photovoltaics,[74–77] and are capable of acting as chiroptical switches.[78–80] Computational studies are often paired with experimental helicene studies to analyze the effect of conformational changes and functional group substitutions, which have a pronounced effects on the electronic properties.[79–89]

The goal of this work is to provide or utilize tools that may lead to predicting the effect of structural or environmental changes on optical rotation. The work in this document is all motivated by these goals, beginning by understanding the physics leading to the generation of natural optical activity (Chapter 2). We investigated helicenes by altering functionalization and geometry (Chapter 3) and rigorously tested a simple solvation model in an attempt to capture solvent shifts or trends

(Chapter 4). We have implemented and tested a method to select orbitals with small contributions to $[\alpha]_\omega$ and eliminate them prior to the calculation (Chapter 5), and have constructed a small basis set with the goal of reproducing the accuracy of larger basis sets at a reduced computational cost (Chapter 6). We have worked in concert with an experimental group to simulate the spectra of the helicenes they synthesize in hopes of one day directing them in the design of chiral molecules with targeted optical properties of interest (Chapter 7).

Chapter 3 discusses our analysis of the origin of the large OR in helicenes, by decomposing this property in terms of one-electron excitations between molecular orbitals. We explore how the OR changes with the size of the helicene, and with the introduction of electron donating and withdrawing groups. We found that the $[\alpha]_\omega$ is dominated by a few transitions with large magnetic dipoles, which can be broken down into three categories. In the same chapter, we also explore the use of localized orbitals, which are often closer to a chemically intuitive picture of single and double bonds. We found that localized orbitals offer a clearer description of the phenomena which lead to optical rotation than the cannonical orbitals; however, the localization scheme is limited by the number of contributions which need to be considered.

In concert with the Vaccaro group at Yale University, who provided us with experimental gas and solution phase data of a series of rigid chiral molecules, we investigated if PCM, a simple solvent model, could reproduce trends in solvent shifts to $[\alpha]_\omega$. This work is presented in Chapter 4, where we analyze the $[\alpha]_\omega$ for each molecule and compare it to experimental values across multiple wavelengths and in multiple solutions of varying polarity. The calculations were investigated both for quantitative reproduction of $[\alpha]_\omega$ in the solvent and gas phases and qualitative agreement in the relative solvent shifts. We found that calculations reproduced gas phase $[\alpha]_\omega$ fairly well; however, PCM could not reproduce the magnitude or ordering of solvent shifts. Additionally, ZPVCs to $[\alpha]_\omega$ were calculated to evaluate their effect on agreement with experimental results. We found that ZPVCs worsened the calculated agreement with experiment when calculations over-predicted $[\alpha]_\omega$; however, under-predicted results were improved.

Chapter 5 recounts our work to develop a method to reduce the cost of $[\alpha]_\omega$ calculations by

reducing the cost of solving the coupled perturbed (CP) Kohn-Sham (KS) or Hartree-Fock (HF) equations. The method selects orbitals based on predicted contributions to $[\alpha]_\omega$ and eliminates them from the calculation. We tested four cutoff types in addition to three cutoff orders of magnitude, two functionals, and two basis sets, and found that the method successfully eliminated un-needed contributions, and a full implementation would lower the overall cost of calculations.

Our work in Chapter 6 is based on the work of Wiberg *et al.*, who showed that combining the diffuse functions of a large basis with a small basis set improved the ability of the small basis set to predict $[\alpha]_\omega$ values similar to the larger basis set.[61] We describe our optimization of the exponents of the diffuse functions for $[\alpha]_\omega$ calculations and then our tests of the reduced basis sets. The new basis sets are smaller in size but able to accurately reproduce the values calculated with a larger basis set.

Chapter 7 details our collaboration with the Avarvari group from the Université d'Angers, who study helicenes due to their unique optical properties and possible applications across a wide variety of organic systems. We simulated electronic spectra in order to investigate the cause of fluorescence quenching they were seeing in their experiments, and successfully reproduced the absorption and emission of their helicenes. We then proceeded to simulate ECD spectra to compare with molecules for which they were able to create enantiomerically pure solutions.

Our work has opened new avenues for the investigation of optical activity. We demonstrated a method to decompose $[\alpha]_\omega$ into component parts, which we used to understand the effect functionalization plays on helicenes, but can be applied to quantum mechanical solvent systems or new molecules of interest in order predict the effect of structural and environmental changes on OR. Reducing the cost of $[\alpha]_\omega$ calculations allows us to access new systems for study or perform calculations of older systems with better descriptions of environmental effects.

Chapter 2

Methods

2.1 Theory of Optical Rotation

2.1.1 The Principles of Optical Rotation

Optical activity, which refers to the different interaction of the two enantiomers of chiral molecules with circularly polarized light, was initially recognized with the work of Arago in 1811, who observed the effect in quartz crystals.[90] One year later, Biot noted two separate effects, the rotation of a plane of linearly polarized light (optical rotation), and then the wavelength dependence to the degree of rotation (optical rotatory dispersion). Several years later, in 1818, he deduced that the angle of rotation was inversely proportional to the square of the wavelength of the incident light.[90] In 1825, Fresnel discovered that plane polarized light could be described as a superposition of left and right circularly polarized light, a discovery that allowed him to relate optical rotation (OR) to a difference in the group velocity of the left and right circularly polarized waves of light in an optically active medium.[90]

This phenomenon can be understood by considering a linearly polarized wave of light with angular frequency ω:

$$\omega = 2\pi c/\lambda \tag{2.1}$$

where c is the speed of light in a vacuum and λ is the wavelength. This wave can be represented as the sum of a left and right circularly polarized waves, each with their own group velocity. If the left and right polarized waves of light enter a medium at the same time, then after some distance l, those individual electric fields are rotated by an angle θ:

$$\theta^R = -2\pi cl/\lambda v^R$$
$$\theta^L = 2\pi cl/\lambda v^L \tag{2.2}$$

where v^R and v^L are the group velocities of the right and left circularly polarized waves, and θ^R and θ^L are the respective angles of rotation. The plane of polarization of the incident light is given by the average of the two angles, θ^R and θ^L, which gives α_λ:

$$\alpha_\lambda = \frac{1}{2}(\theta^R + \theta^L) = \frac{\pi l}{\lambda}(n^L - n^R) \tag{2.3}$$

where n^L and n^R are the left and right indices of refraction, given by $n = c/v$. Therefore, the angle of polarization of the transmitted light, and thus the optical rotation, are dependent on the circular birefringence of the medium, which is the difference in index of refraction for left and right circularly polarized light. The optical rotation is usually reported normalized to the path length and concentration, denoted as the specific rotation at a certain wavelength λ:

$$[\alpha]_\lambda = \frac{\alpha_\lambda}{M * l} \tag{2.4}$$

2.1.2 The Quantum Mechanical Description of Optical Rotation

The derivation of the quantum mechanical formula of the specific rotation begins from considering a chiral system in the presence of an electromagnetic field. We use perturbation theory, starting with a free particle for simplicity, and reverting back to molecules later on. A free charged particle in an electromagnetic field experiences a Lorentz force:[90]

$$\mathbf{F} = e\left[\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B}\right] \tag{2.5}$$

where $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic fields, respectively, and e is the charge of an electron moving with velocity $\mathbf{v}$. We find that the following Lagrangian

$$L = \frac{1}{2}mv^2 - e\phi + \frac{e}{c}\mathbf{A}\cdot\mathbf{v} \tag{2.6}$$

where ϕ and $\mathbf{A}$ are the scalar and vector potentials, m is the mass of the particle, $\mathbf{v}$ the velocity, and e the charge,[90] satisfies the the Euler-Lagrange equation:

$$\nabla L - \frac{\mathrm{d}}{\mathrm{dt}}\frac{\partial L}{\partial \mathbf{v}} = 0 \tag{2.7}$$

From equation 2.6, we find the Hamiltonian using

$$\begin{aligned} H &= \mathbf{v}\cdot\frac{\partial L}{\partial \mathbf{v}} - L \\ &= \frac{1}{2m}(\mathbf{p} - e\mathbf{A})^2 + e\phi \end{aligned} \tag{2.8}$$

where $\mathbf{p}$ is the generalized momentum of the particle. This Hamiltonian may be subdivided into reference and perturbative parts, $H = H_0 + H'$:

$$H_0 = \frac{p^2}{2m} \tag{2.9}$$

$$H' = -\frac{e}{m}\mathbf{p}\cdot\mathbf{A} + e\phi \tag{2.10}$$

where the term in A^2 is not included, as it does not contribute in a semi-classical treatment of light scattering. As the electric and magnetic fields are determined using

$$\begin{aligned} \mathbf{E} &= -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} \\ \mathbf{B} &= \nabla\times\mathbf{A} \end{aligned} \tag{2.11}$$

we are left to choose the specific form of the scalar and vector potentials. We expand about the center of the particle in a translating frame and make the choice:[90]

$$\begin{aligned} \phi(\mathbf{r}) &= (\phi)_0 - r_\alpha(E_\alpha)_0 + ... \\ A_\alpha(\mathbf{r}) &= \frac{1}{2}\varepsilon_{\alpha\beta\gamma}(B_\beta)_0 r_\gamma + ... \end{aligned} \tag{2.12}$$

because higher order multipole moments are required to describe the effects of a non-static field. Substituting Eq. 2.12 into Eq. 2.10, we arrive at an expression for the interaction Hamiltonian:[90]

$$H' = e(\phi)_0 - \mu_\alpha(E_\alpha)_0 - m_\alpha(B_\alpha)_0 + ... \tag{2.13}$$

We switch now to a description in terms of a molecule in an isotropic medium, rather than of a single free particle. In order to obtain the optical rotation, we must find the molecular property tensors in an oscillating field. The perturbed molecular wavefunction may be obtained from the time-dependent Schrödinger equation:

$$\left(i\hbar\frac{\partial}{\partial t} - H_0\right)\psi = H'\psi \tag{2.14}$$

where H_0 is now the unperturbed molecular Hamiltonian and H' is the interaction Hamiltonian from Eq. 2.13. We begin from the stationary state solution to the time-dependent equation

$$\psi_n = \psi_n^{(0)} e^{-i\omega_n t} \tag{2.15}$$

where $\psi_n^{(0)}$ is the solution to the time-independent Schrödinger equation. Treating the fields as a harmonic perturbation of frequency ω, and assuming that the wavelength of incident light is much larger than the size of the molecule, we may develop a form for the perturbed eigenfunction which is first order in the perturbations.[90] We may then write the induced multipole moments of the molecule using the expectation value of the corresponding operator. Rosenfeld showed that the electric dipole of a chiral molecule induced by a frequency-dependent electromagnetic field can be written as:[91]

$$\langle\hat{\boldsymbol{\mu}}\rangle = \boldsymbol{\mu}_0 + \alpha\cdot\mathbf{E} + \frac{1}{\omega}G'\frac{\partial\mathbf{B}}{\partial t} \tag{2.16}$$

where $\langle\hat{\boldsymbol{\mu}}\rangle$ is the expectation value of the electric dipole operator, $\boldsymbol{\mu}_0$ is the static dipole moment of the molecule, α is the polarizability tensor, and G' is the mixed electric dipole-magnetic dipole

polarizability tensor:

$$\begin{aligned} G'(\omega) &= -\frac{1}{\hbar}\sum_{j\neq 0}\mathrm{Im}\left[\frac{\langle\psi_0|\hat{\boldsymbol{\mu}}|\psi_j\rangle\langle\psi_j|\hat{\mathbf{m}}|\psi_0\rangle}{\omega_{j0}-\omega}+C.C.\right] \\ &= -\frac{2\omega}{\hbar}\sum_{j\neq 0}\frac{\mathrm{Im}(\langle\psi_0|\hat{\boldsymbol{\mu}}|\psi_j\rangle\langle\psi_j|\hat{\mathbf{m}}|\psi_0\rangle)}{\omega_{j0}^2-\omega^2} \end{aligned} \tag{2.17}$$

where ψ_0 is the ground electronic state of the molecule, ψ_j is the jth electronic excited state, ω_{j0} is the energy difference between the two states, and $\hbar$ is the reduced Planck's constant. The G' tensor is related to the specific rotation of a molecule in an isotropic medium:

$$[\alpha]_\omega = \frac{(72.0\times 10^6)\hbar^2 N_A\omega}{c^2 m_e^2 M}\times\left[\frac{1}{3}\mathrm{Tr}(G')\right] \tag{2.18}$$

where N_A is Avogadro's number, m_e is the mass of the electron, and M is the molecular mass in amu.[16] The specific rotation in Eq. 2.18 is the same as that in Eq. 2.4, but expressed in terms of frequency rather than wavelength. We note that to perform the calculation for an oriented system, the contribution from the mixed electric dipole-electric quadrupole polarizability tensor, $A_{\alpha,\beta\gamma}$, is needed. As this tensor is traceless, it does not contribute in isotropic systems.

2.2 Linear Response

The definition of G' in Eq. 2.17 is based on a sum over the excited states of the molecule. In practice, the sum-over-states series converges very slowly,[92] and linear response (LR) techniques are used instead.[93] The LR expression can be recovered by inserting the resolution of the identity, $\sum_k|\psi_k\rangle\langle\psi_k| = 1$, in Eq. 2.17 to obtain:

$$G'(\omega) = -\frac{1}{\hbar}\sum_{j\neq 0}\sum_k \mathrm{Im}[\langle\psi_0|\hat{\boldsymbol{\mu}}|\psi_j\rangle\langle\psi_j|(\omega_{j0}-\omega)^{-1}|\psi_k\rangle\langle\psi_k|\hat{\mathbf{m}}|\psi_0\rangle + C.C.] \tag{2.19}$$

collecting terms, we may write:

$$\langle \psi_j | \mathbf{X}_m(\pm\omega) | \psi_0 \rangle = \sum_k \langle \psi_j | (\omega_{j0} \mp \omega)^{-1} | \psi_k \rangle \langle \psi_k | \hat{\mathbf{m}} | \psi_0 \rangle \tag{2.20}$$

where $\langle \psi_j | \mathbf{X}_m(\pm\omega) | \psi_0 \rangle$ is the perturbed density of the molecule under the effect of a magnetic field. By solving Eq. 2.20 in the space of excited Slater determinants $\{\Psi_j\}$ (built from the ground state Ψ_0 as a series of molecular orbital excitations) instead of excited states, and inverting the matrix on the right-hand side, we obtain a system of linear equations which we can solve for $\mathbf{X}_m(\pm\omega)$:

$$\sum_{l \neq 0} \langle \Psi_m | (H - E_0 \mp \omega | \Psi_l \rangle \langle \Psi_l | \mathbf{X}_m(\pm\omega) | \Psi_0 \rangle = \langle \Psi_m | \hat{\mathbf{m}} | \Psi_0 \rangle \tag{2.21}$$

where $E_0 = \langle \Psi_0 | H | \Psi_0 \rangle$ is the energy of the unperturbed molecule. This method is more efficient than the sum-over-states approach, because instead of having to evaluate all the excited states, we need only solve the system of linear equations in Eq. 2.21, and then insert the solution into Eq. 2.19.

2.3 Choice of Gauge

Practical calculations of $[\alpha]_\omega$ are dependent on the choice of molecular origin. If the origin of the coordinate system is shifted by some vector, $O' = O + \mathbf{a}$, the numerator of G':

$$\omega \mathrm{Im}\Big(\sum_i e_i \langle \Psi_0 | \hat{\mathbf{r}}_i | \Psi_j \rangle \cdot \sum_k \frac{e_k}{2m_k} \langle \Psi_j | \hat{\mathbf{r}}_k \times \hat{\mathbf{p}}_k | \Psi_0 \rangle \Big) \tag{2.22}$$

shifts to:

$$\omega \mathrm{Im}\Big(\sum_i e_i \langle \Psi_0 | (\hat{\mathbf{r}}_i - \mathbf{a}) | \Psi_j \rangle \cdot \sum_k \frac{e_k}{2m_k} \langle \Psi_j | (\hat{\mathbf{r}}_k - \mathbf{a}) \times \hat{\mathbf{p}}_j | \Psi_0 \rangle \Big) \tag{2.23}$$

in the length gauge (LG), where the sums are over the number of electrons in the molecule, and we have used the definitions $\hat{\boldsymbol{\mu}} = \sum_i e_i \mathbf{r}_i$ and $\hat{\mathbf{m}} = \sum_i \frac{e_i}{2m_i} \mathbf{r}_i \times \mathbf{p}_i$. In the following equations, the leading

ω, constants, sums, and subscripts will be dropped for clarity. Equation 2.23 is expanded to:

$$\begin{aligned}&\langle\Psi_0|\hat{\mathbf{r}}|\Psi_j\rangle\cdot\langle\Psi_j|\hat{\mathbf{r}}\times\hat{\mathbf{p}}|\Psi_0\rangle-\langle\Psi_0|\mathbf{a}|\Psi_j\rangle\cdot\langle\Psi_j|\hat{\mathbf{r}}\times\hat{\mathbf{p}}|\Psi_0\rangle\\&-\langle\Psi_0|\hat{\mathbf{r}}|\Psi_j\rangle\cdot\langle\Psi_j|\mathbf{a}\times\hat{\mathbf{p}}|\Psi_0\rangle+\langle\Psi_0|\mathbf{a}|\Psi_j\rangle\cdot\langle\Psi_j|\mathbf{a}\times\hat{\mathbf{p}}|\Psi_0\rangle\end{aligned}\tag{2.24}$$

Because the functions Ψ_j are orthogonal to Ψ_0, the second and fourth terms in Eq. 2.24 naturally disappear ($\langle\Psi_0|\mathbf{a}|\Psi_j\rangle=\mathbf{a}\langle\Psi_0|\Psi_j\rangle=0$), and the origin dependence remains only in the third term. In the limit of a complete basis set, we can use the relation $-i\frac{\hbar}{m}\hat{\mathbf{p}}=[\hat{\mathbf{r}},\hat{H}]$ in Eq. 2.24:

$$\langle\Psi_0|\hat{\mathbf{r}}|\Psi_j\rangle\cdot\langle\Psi_j|\hat{\mathbf{r}}\times\hat{\mathbf{p}}|\Psi_0\rangle-\langle\Psi_0|\hat{\mathbf{r}}|\Psi_j\rangle\cdot\langle\Psi_j|\mathbf{a}\times[\hat{\mathbf{r}},\hat{H}]|\Psi_0\rangle\tag{2.25}$$

Because $\hat{H}$ is a Hermitian operator, we can write the second term in equation 2.25 as:

$$\omega_{j0}\langle\Psi_0|\hat{\mathbf{r}}|\Psi_j\rangle\cdot\langle\Psi_j|\mathbf{a}\times\hat{\mathbf{r}}|\Psi_0\rangle\tag{2.26}$$

which is now 0 due to the properties of a dot and cross product.

In practical calculations, the use of a complete basis set is not possible and the problem of origin dependence remains. Modern quantum calculations alleviate this issue by using London atomic orbitals, also known as gauge invariant atomic orbitals (GIAOs).[19,94,95] GIAOs add a field dependent phase factor to a standard atomic orbital χ_ν:

$$\xi_\nu(t)=\exp\left\{-\frac{i}{2}[(\mathbf{B}^+e^{i\omega t}+\mathbf{B}^-e^{-i\omega t})\times\mathbf{R}_\nu]\cdot\mathbf{r}\right\}\chi_\nu\tag{2.27}$$

where $\mathbf{R}_\nu$ is the coordinate of the center of the orbital. After some algebra,[96,97] GIAOs lead to the same replacement as in equation 2.25, and the origin dependence of the magnetic dipole once again disappears.

Another solution to the gauge problem is to perform $[\alpha]_\omega$ calculations in the velocity gauge, where Eq. 2.22 becomes:

$$\frac{1}{\omega}\langle\Psi_0|\hat{\mathbf{p}}|\Psi_j\rangle\langle\Psi_j|\hat{\mathbf{r}}\times\hat{\mathbf{p}}|\Psi_0\rangle\tag{2.28}$$

This naturally removes the issue of a shift in origin, although a new problem arises. The $\frac{1}{\omega}$ prefactor cancels with ω in the numerator of Eq. 2.18, causing $[\alpha]_\omega$ to have an unphysical non-zero value at the static limit (if $\omega = 0$, there is no light and there should therefore be no plane of polarization to rotate). One solution, proposed by Pederson *et al.* and called the modified velocity gauge (MVG), is to subtract out the static limit from the calculated value of $[\alpha]_\omega$.[17] At the complete basis set limit, the static limit term goes to 0 and the LG and MVG calculations become equivalent. However, because $[\alpha]_\omega$ calculated in the LG converges faster to the complete basis set limit with the size of the basis set than in the MVG, LG calculations are preferred when using DFT.[17] On the other hand, because coupled cluster (CC) calculations ignore the response of the orbitals, GIAOs do not solve the origin dependence in the LG. Ruud *et al.* studied the effect of moving the origin in CC calculations, and suggested either to place the origin at the center of mass of the molecule, or to use the MVG.[13]

Chapter 3

Configuration Space Analysis of the Specific Rotation of Helicenes

(This work taken with the permission of Tal Aharon and Marco Caricato from *J. Phys. Chem. A* **2019**, *123*, 4406-4418.[98] Supporting information is available online.)

3.1 Introduction

The study of optically active molecules is of continuous interest due to the fundamental role chiral molecules play in living organisms, as building blocks of life (i.e., L-amino acids and D-sugars) as well as drugs.[2] Given the homochiral nature of biological compounds, the correct determination of the absolute configuration is paramount, for instance, for the synthesis of natural products with favorable biological activity. Chiroptical spectroscopy and theoretical simulations have been very successful at this task, and they have become an essential tool in the pharmaceutical industry. This work focuses on the oldest of these measurements, that of the specific rotation $[\alpha]_\omega$, whose sign is directly related to the absolute configuration of a particular enantiomer. The accurate calculation of $[\alpha]_\omega$ is difficult as this property is very sensitive to the molecular structure and the interaction with the environment. Nevertheless, great progress has been made over the past two decades, including the use of modern density functional theory (DFT) and coupled cluster (CC) methods.[5–20]

Despite this progress, a chemically intuitive understanding of the relationship between the absolute configuration of a chiral molecule and the $[\alpha]_\omega$ sign and magnitude is still lacking. Various groups have developed models to determine structure-property relationships. For instance, Autschbach and coworkers performed a decomposition of $[\alpha]_\omega$ via localized orbital contributions

to the diagonal elements of the Rosenfeld tensor. This work showed that the large $[\alpha]_\omega$ of (1S,4S)-norbornenone is due to electron delocalization between the C=O and C=C chromophores.[21] Kahr and coworkers extended the analysis of optical activity to include non-chiral molecules through a simple and effective Hückel theory decomposition.[24,99] Wiberg examined the effect on $[\alpha]_\omega$ of the rotation of the torsional angle of terminally substituted 1,4-pentadiene (C=C to C=X interactions, where X=C, O, NH, and S), and found a dependence on the electronegativity of X.[100,101] We have also introduced a method of decomposing $[\alpha]_\omega$ in terms of transition electric and magnetic dipole contributions, called $\tilde{S}_k$ method (discussed in section 3.2).[23,26] We showed that a limited number of orbital transitions can be used to describe $[\alpha]_\omega$ for molecules with strong chromophoric groups,[23] and that changes in $[\alpha]_\omega$ due to conformational flexibility are dominated by a small number of transitions,[26] thus simplifying the interpretation of the structure-property relationship.

In this work, we apply the $\tilde{S}_k$ analysis to a series of standard and functionalized helicenes of various length. These molecules are comprised of a series of fused phenyl rings, and have been studied extensively both computationally and experimentally.[67] They are known for their intense chiroptical response,[102] non-linear optical activity,[103,104] and self-assembly.[67] They have also garnered interest as redox active chiral switches,[79,83] and have been functionalized extensively to modulate their optical properties.[81] The Avarvari group created many helicene derivatives to investigate the effect of functionalization on its electronic properties.[81] Crassous and coworkers successfully synthesized novel organometallic helicenes,[81,87,105–107] and showed that the electronic circular dichroism (ECD) spectrum reversibly switches under redox stimuli.[79,83] Autschbach and coworkers performed theoretical simulations of ECD and UV/Vis absorption spectra of these organometallic helicenes to elucidate the role of charge-transfer excitations in the redox switching mechanism.[82,84–86,88] Nakai, Mori, and Inoue also studied the ECD spectra of helicenes, and analyzed the rotatory strength of various absorption bands based on the relative orientation of the transition electric and magnetic dipole moments.[108–110] In this work, we extend these analyses to the specific rotation of standard and functionalized helicenes through the $\tilde{S}_k$ method. We aim to understand how $[\alpha]_\omega$ changes with length and pitch of the helical backbone, and with the introduction

of electron-withdrawing or donating groups, which may lead to guidelines for the development of compounds with desired electronic responses. The $\tilde{S}_k$ analysis reveals that the $[\alpha]_\omega$ magnitude is driven by delocalized motion of the electron along the entire backbone of the helicene, resulting in strong magnetic dipole contributions. Functionalization directs the orbital density towards or away from the substituent groups, affecting $[\alpha]_\omega$ based on the electron push/pull strength and on the degree of orbital overlap between the groups.

Since $[\alpha]_\omega$ is invariant to unitary transformations within the occupied and virtual molecular orbital (MO) subspaces, we also explore the use of localized MOs (LMOs) (with the Boys[111,112] and Pipek-Mezey[113] approaches) for the $\tilde{S}_k$ decomposition of the specific rotation. We find that LMOs may provide a more chemically intuitive picture than canonical MOs, however, they are best employed when the molecule possesses small, localized chromophores.

This work is organized as follows: section 3.2 reviews the $\tilde{S}_k$ method and the derivation of localized MOs, section 3.3 describes the computational protocol, section 3.4 presents the results of the calculations, and section 3.5 contains a discussion of these results and concluding remarks.

3.2 Theory

From Rosenfeld's derivation, the specific rotation $[\alpha]_\omega$ can be expressed using a sum-over-states (SOS) formula as:

$$[\alpha]_\omega = -\frac{C_\omega}{M}\sum_{n=1}^{N_{ex}}\left(\frac{R_n}{\omega_n-\omega}+\frac{R_n^*}{\omega_n+\omega}\right) \tag{3.1}$$

where N_{ex} is the number of excited states, ω_n is the transition energy to the nth excited state, ω is the frequency of the incident light, M is the molecular mass, and C_ω is a proportionality constant to obtain the common units of deg [dm (g/mL)]$^{-1}$.[91] The rotatory strength R_n is related to the differential absorption of left and right circularly polarized light measured in circular dichroism spectroscopy:

$$R_n = \mathrm{Im}\left\{\sum_{i=1}^{3}\langle 0|\mu_i|\Psi_n\rangle\langle\Psi_n|m_i|0\rangle\right\} \tag{3.2}$$

where μ_i and m_i are the ith component of the electric and magnetic dipole operators, respectively, and Ψ_n is the nth excited state. Note that R_n is the dot product of the two transition dipoles. If the energy denominator were to be included into the definition of R_n, Eq. 3.1 may be written in a more compact form:

$$[\alpha]_\omega = -\frac{C_\omega}{M}\sum_{n=1}^{N_{ex}}\left(\tilde{R}_n + \tilde{R}_n^*\right) \tag{3.3}$$

Eqs. 3.1-3.3 apply only in nonresonant conditions (i.e., ω far from ω_n), as is the case in this work.

The SOS equations are of little practical use, due to the large number of excited states needed to converge the series in Eq. 3.3.[92] A more computationally efficient approach to calculate $[\alpha]_\omega$ is based on the linear response formalism.[15,20] For the case of single reference methods such as Hartree-Fock (HF) or Kohn-Sham density functional theory (KS DFT), $[\alpha]_\omega$ becomes

$$[\alpha]_\omega = -\frac{C_\omega}{M}\mathrm{Im}\left\{\sum_{i=1}^{3}\sum_{k=1}^{N_{ex}}\left(\langle 0|\mu_i|\phi_k\rangle\langle\phi_k|X_{m_i}^+|0\rangle + \mathrm{c.c}\right)\right\} \tag{3.4}$$

where k runs over singly excited Slater determinants ϕ_k. The $\langle\phi_k|X_{m_i}^+|0\rangle$ matrix element represents the electron density perturbed by the magnetic dipole, and it is evaluated by solving the coupled-perturbed (CP) HF or KS equations:

$$\langle\phi_k|X_{m_i}^\pm|0\rangle = \sum_{l=1}^{N_{ex}}\langle\phi_k|(\hat{H} - E_0 \mp \omega)|\phi_l\rangle^{-1}\langle\phi_l|m_i|0\rangle \tag{3.5}$$

In order to interpret the specific rotation in terms of molecular orbital (MO) contributions, we may define a quantity similar to $\tilde{R}_n$ in the configuration space of the Slater determinants by switching the order of the sums in Eq. 3.4:

$$\tilde{S}_k = \mathrm{Im}\left\{\sum_{i=1}^{3}\langle 0|\mu_i|\phi_k\rangle\langle\phi_k|X_{m_i}^+|0\rangle\right\} \tag{3.6}$$

so that Eq. 3.4 can be re-written as a sum of $\tilde{S}_k$ rotatory strengths:

$$[\alpha]_\omega = -\frac{C_\omega}{M}\sum_{k=1}^{N_{ex}}\left(\tilde{S}_k + \tilde{S}_k^*\right) \tag{3.7}$$

Since $\tilde{S}_k$ is also written as a dot product between two vectors, its magnitude and sign depend on the magnitude of the vectors and the angle θ between them. The advantage of the linear response formalism over the SOS formula is that the Slater determinant contributions to $[\alpha]_\omega$ (and thus the $\tilde{S}_k$ values) can be computed all at once by solving the linear system of equations in Eq. 3.5.

In section 3.4, we proceed to analyze the origin of $[\alpha]_\omega$ of each molecule by the components of the $\tilde{S}_k$ dot product as well as the involved occupied and virtual MOs with Eqs. 3.6-3.7. One can focus on the largest $\tilde{S}_k$ contributions to the specific rotation to obtain a qualitative understanding of how individual orbital transitions affect this property, which may provide key insights on the structure-property relations in chiroptical spectroscopy. For clarity of discussion, in the rest of the paper we refer to "$\tilde{S}_k$ contributions" to $[\alpha]_\omega$ even if the values reported are actually $\tilde{S}_k + \tilde{S}_k^*$ as in Eq. 3.7, and we use "magnetic dipole" contributions when discussing the $\langle\phi_k|X_{m_i}^+|0\rangle$ matrix elements in Eq. 3.6.

The Rosenfeld tensor and $[\alpha]_\omega$ are invariant under unitary transformations within the occupied MO and virtual MO subspaces, respectively. Therefore, it may be convenient to express $\tilde{S}_k$ using different choices of MO basis, beyond that of canonical MOs, as different bases may provide a smaller number of $\tilde{S}_k$ contributions to $[\alpha]_\omega$, or a more chemically intuitive picture. Localized MOs (LMOs) are often utilized for this purpose. Therefore, we consider two popular choices of LMOs: Boys[111,112] and Pipek-Mezey (PM)[113]. The former are obtained by maximizing the distance between orbital centroids:

$$\sum_{j=1}^{N_{MO}}\sum_{i>j}^{N_{MO}}\left[\langle i|\mathbf{r}|i\rangle - \langle j|\mathbf{r}|j\rangle\right]^2 \tag{3.8}$$

where N_{MO} is the number of MOs in the subspace. The PM localization is based on the definition

of a localization parameter d_i, which counts how many atoms are spanned by the ith MO:

$$d_i = \left[\sum_{A=1}^{N_{at}} (Q_A^i)^2 \right]^{-1} \tag{3.9}$$

where N_{at} is the number of atoms in the molecule, and Q_A^i is the Mulliken population of orbital i on atom A. From d_i, one can define a mean delocalization of all MOs in the subspace:

$$D = N_{MO} \left(\sum_{i=1}^{N_{MO}} d_i^{-1} \right)^{-1} \tag{3.10}$$

and the LMOs are obtained by minimization of D. Both localization techniques scale as $O(N^3)$, where N is the atomic basis set size, as they require the evaluation of electric dipole integrals. Thus, the computational cost of the MO localization is small compared to the solution of the self-consistent field (SCF) and CP equations.

3.3 Computational Details

All calculations were performed with a development version of the GAUSSIAN suite of programs.[114] Geometries were optimized using the CAM-B3LYP/aug-cc-pVDZ model chemistry with D3 dispersion corrections.[115,116] Calculations of specific rotation were performed at the sodium-D line ($\omega = 589.3$ nm) with the B3LYP[117–119]/p-3-21G level of theory in the length gauge formalism using gauge-including atomic orbitals[19,94,95]. The p-3-21G basis set is built by augmenting the standard 3-21G basis set with p-type diffuse functions from the standard aug-cc-pVDZ basis set. This unorthodox choice is based on previous work that showed how diffuse functions are the most important to obtain a good description of $[\alpha]_D$.[61] In the context of this study, using a reduced basis set is critical as this allows to reduce the number of $\tilde{S}_k$ contributions, and to simplify the qualitative decomposition and analysis of $[\alpha]_D$ in terms of orbital transitions. Nevertheless, all $[\alpha]_D$ values computed with p-3-21G are in reasonable agreement with those obtained with the full aug-cc-pVDZ basis set (with errors of the order of 9%-18%), thus justifying the use of this smaller

set.

3.4 Results

3.4.1 Localized Orbitals

As discussed in section 3.2, the Rosenfeld tensor (and thus $[\alpha]_\omega$) can be decomposed in different MO bases, related by unitary transformations within the occupied MO or the virtual MO subspaces. Thus, we consider transformations of both the occupied and virtual MO subspaces, as well as transformation of only the occupied MO subspace, as this may be enough for the qualitative interpretation of the $[\alpha]_D$ tensor with the $\tilde{S}_k$ analysis. In summary, we obtain the $[\alpha]_\omega$ tensor and the corresponding $\tilde{S}_k$ values for five choices of MO basis: canonical MOs (CMOs), Boys-localized occupied MOs (BO-LMOs), Boys-localized occupied and virtual MOs (BOV-LMOs), PM-localized occupied MOs (PMO-LMOs), and PM-localized occupied and virtual MOs (PMOV-LMOs). Given that the computational cost of the MO localization and of the tensor transformation is virtually negligible compared to the solution of the SCF and linear response equations, we evaluate $[\alpha]_D$ and $\tilde{S}_k$ with all five MO options for every calculation, and choose the most appropriate for a specific system on-the-fly.

The first test case is (1S,4S)-norbornenone, because this molecule has a very large $[\alpha]_D$, and previous studies have shown that it is dominated by a small number of electronic transitions.[21–23] In Figure 3.1, we report the cumulative contribution of the $\tilde{S}_k$ transitions to the total $[\alpha]_D$, separated according to their magnitude. The canonical MO basis provides the most compact representation of $[\alpha]_D$, in the sense that a single transition is dominant, and it determines the sign. More transitions are necessary to reach the converged value of the rotation (*e.g.*, 8 transitions are necessary to be within 50% of the converged value, and about 60 transitions to be within the 10% margin), but a single transition is enough for a qualitative discussion. The isodensity plots and corresponding vector representation of the electric and magnetic dipoles for this transition are reported in Figure S1 of the supporting information (SI). Since the interpretation of this transition was already

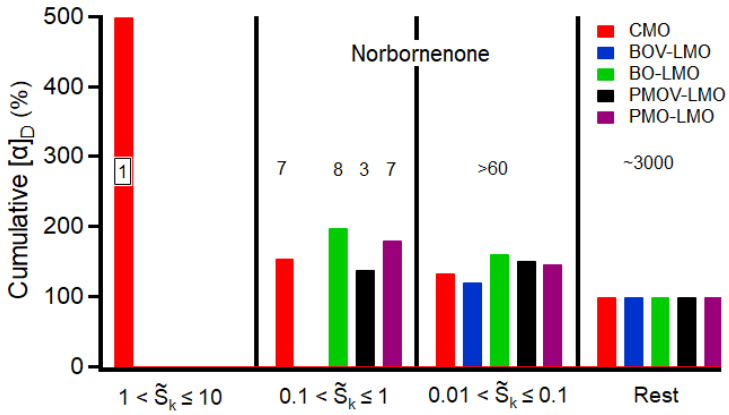
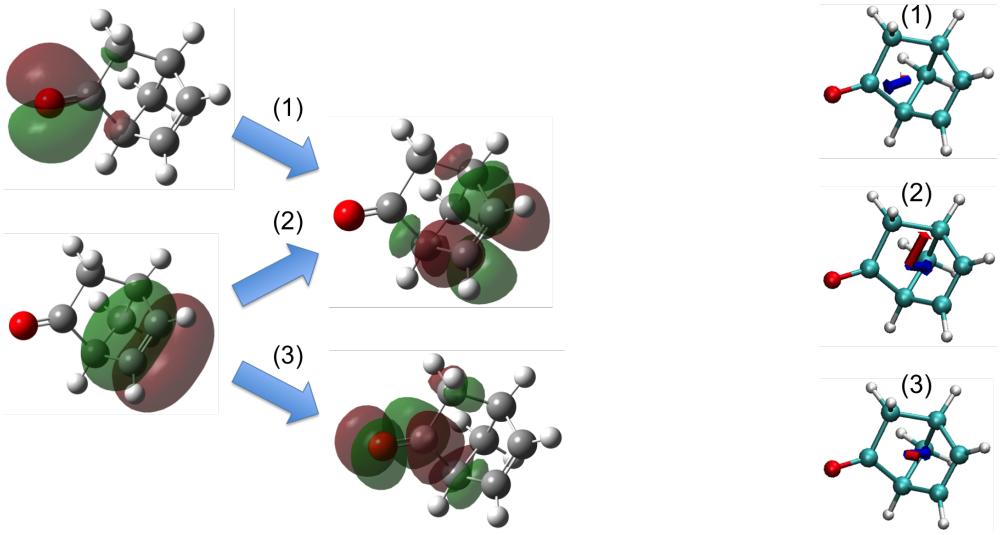


Figure 3.1: Plot of the cumulative % $\tilde{S}_k$ contributions to the total $[\alpha]_D$ for (1S,4S)-norbornenone with various choices of MO basis. The $\tilde{S}_k$ contributions are divided according to their magnitude as reported in the x axis labels, while the y axis reports the cumulative percent of $[\alpha]_D$ recovered by summing all $\tilde{S}_k$ in the current and larger-value ranges. The numbers on the bars indicate how many transitions fall within each range.

presented in our previous work,[23] we do not further discuss it here. In terms of LMO bases, the PMOV-LMO choice also provides a fairly compact representation of the specific rotation, as only three transitions are sufficient to be within 30% of the converged value, see Figure 3.1. The iso-density plots for the four LMOs involved in the transitions, and the electric and magnetic dipoles are shown in Figure 3.2. Unless explicitly noted, the electric and magnetic dipoles shown in this and in the following figures are unscaled. As expected, the transitions involve orbitals localized on the carbonyl and alkene groups. All $\tilde{S}_k$ values have the same sign, as the angle between the vectors is $< 90°$. Transitions (1) and (3) indicate a charge transfer from the carbonyl to the alkene group and vice versa, with a rotation of the orbital density along the carbonyl group; transition (2) is of $\pi \rightarrow \pi^*$ nature and it would have no contribution to $[\alpha]_D$ in a symmetric alkene group (e.g., ethylene) where the two vectors would be perfectly orthogonal to each other, but it presents a non-negligible $\tilde{S}_k$ value in this case because of the geometrical distortion introduced by the cage structure. The other choices of orbital localization do not offer a similarly compact representation of the tensor, as at least 7-8 transitions need to be considered at once.

We repeat a similar analysis for [8]helicene, with the $\tilde{S}_k$ contributions to $[\alpha]_D$ reported in Figure 3.3. The CMO basis provides the most compact representation of the specific rotation, with 16



(a) Localized MOs (b) Transition dipoles.

Figure 3.2: Transitions with the largest $\tilde{S}_k$ values for (1S,4S)-norbornenone using PMOV-LMOs. Left: occupied and virtual LMOs; right: electric (red) and magnetic (blue) transition dipole moments.

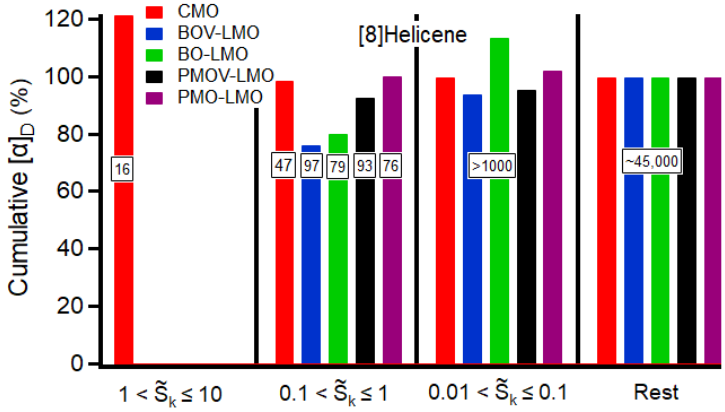


Figure 3.3: Plot of the cumulative % $\tilde{S}_k$ contributions to the total $[\alpha]_D$ for [8]helicene with various choices of MO basis. The $\tilde{S}_k$ contributions are divided according to their magnitude as reported in the x axis labels, while the y axis reports the cumulative percent of $[\alpha]_D$ recovered by summing all $\tilde{S}_k$ in the current and larger-value ranges. The numbers on the bars indicate how many transitions fall within each range.

transitions necessary to get within 20% of the converged $[\alpha]_D$ value. As described more in detail in the next section, these orbitals are delocalized over a large part of the molecule, and not all of the 16 $\tilde{S}_k$ values need to be discussed individually as they correspond to transitions similar in nature. On the other hand, none of the LMO bases provides a compact decomposition of the specific rotation, as at least 76 $\tilde{S}_k$ values are necessary to reproduce $[\alpha]_D$ (with PMO-LMO). Such large number would make the analysis rather cumbersome even if the individual MOs are localized. This result is not completely surprising, as delocalized CMOs can describe rather complex orbital density motion across the molecular frame all at once. This complex motion is fragmented into many contributions using a localized MO picture, resulting in many $\tilde{S}_k$ values of comparable magnitude that need to be analyzed explicitly. This trend holds for all of the other helicene compounds considered in this study, as shown in Figures S2-S6 in the SI, thus, we will continue the discussion only in terms of the canonical MO results. Additional plots detailing the $\tilde{S}_k$ contributions to $[\alpha]_D$ in the CMO basis (divided by $\tilde{S}_k$ magnitude) for all compounds are reported in Figures S7-S9 of the SI.

3.4.2 Helicenes

Helicenes possess strong optical activity, leading to large $[\alpha]_D$, which is clearly due to the intrinsically chiral helical structure of these compounds. However, it is interesting to understand whether this strong chiral response is primarily due to the electric or magnetic dipole components, or their relative orientation. At the same time, it is important to determine how the optical response changes with functionalization of these molecules, as this can provide useful insight for the design of compounds with targeted electronic characteristics. Thus, we consider unfunctionalized helicenes (in section 3.4.2.1) as well as their functionalized equivalent with both electron withdrawing: benzothiadiazole (BTD), and donating groups: two thiol groups (in sections 3.4.2.2 and 3.4.2.3). For the functionalized species, we consider both mono- and bis-substitution. We compare molecules of similar molecular mass, since $[\alpha]_D$ is inversely proportional to this quantity (see Table S1). Thus, we consider the dithiol group as equivalent to a phenyl ring unit, and a BTD group

as equivalent to two phenyl ring units. In other words, we compare the results for [N]helicene (where N is the number of phenyl units) with those for dithiol[N-1]helicene, BTD[N-2]helicene, bis-dithiol[N-2]helicene, bis-BTD[N-4]helicene, and dithiol-BTD[N-3]helicene. The structure of the compounds for N = 6 is shown in Figure 3.4, and the calculated $[\alpha]_D$ values for all helicenes are shown in Figure 3.5.

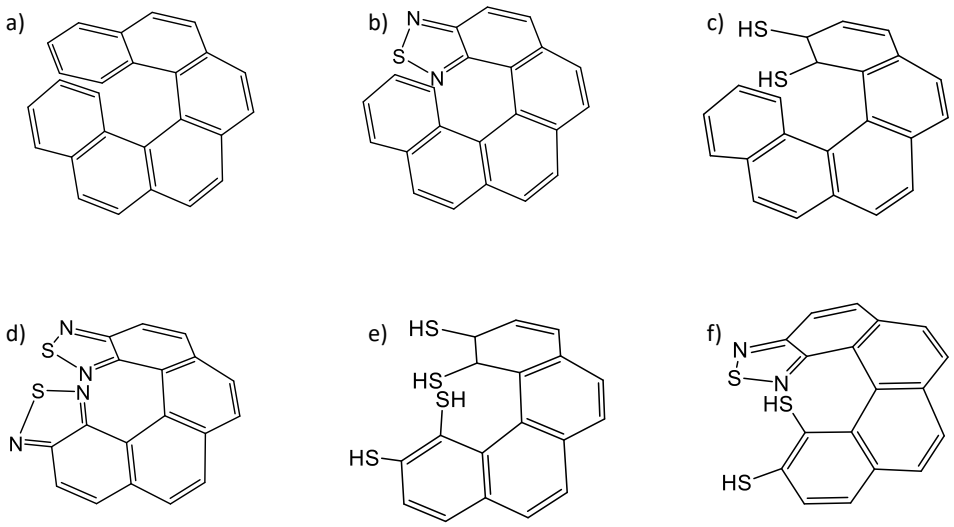


Figure 3.4: Structures of the N = 6 forms of the various helicenes discussed in this work. a) [6]helicene, b) BTD[4]helicene, c) dithiol[5]helicene, d) bis-BTD[2]helicene, e) bis-dithiol[4]helicene, f) dithiol-BTD[3]helicene.

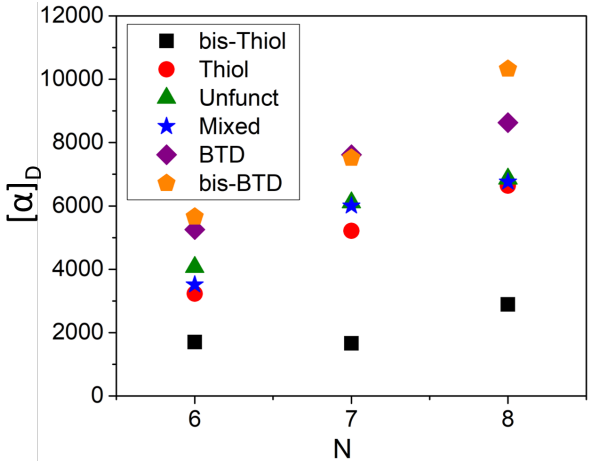
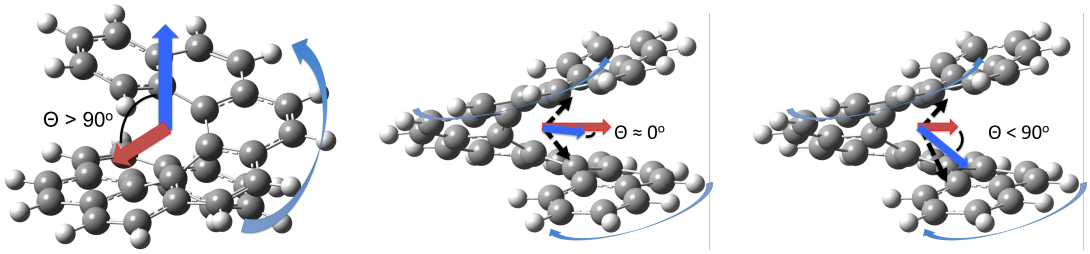


Figure 3.5: Calculated $[\alpha]_D$ (deg [dm (g/mL)]$^{-1}$) for the helicenes in this study. Unfunct: [N]helicene; Thiol: dithiol[N-1]helicene; BTD: BTD[N-2]helicene; Bis-Thiol: bis-dithiol[N-2]helicene; Bis-BTD: bis-BTD[N-4]helicene; Mixed: dithiol-BTD[N-3]helicene.

Using the $\tilde{S}_k$ analysis, we can separate the contributions to $[\alpha]_D$ in two categories: one where the magnetic vector is aligned with the helical main axis, Type A, and one where it is not, Type B; the latter can be further separated in two subcategories: one where the magnetic vector is orthogonal to the helical main axis, Type B_s, and one where the magnetic vector is tilted with respect to the helical main axis (usually by an angle $0^\circ < \phi < 90^\circ$), Type B_a. The *s* subscript stands for *symmetric* because these transitions are significant in molecules with the same terminal groups (and the magnetic vector coincides with the C_2 rotational axis), whereas the *a* subscript stands for *asymmetric* because these transition are significant for compounds with different terminal groups. The transition types are shown schematically in Figure 3.6.



(a) Type A (b) Type B_s (c) Type B_a

Figure 3.6: Types of transitions that dominate the contribution to $[\alpha]_D$. Curved arrows represent the motion of the orbital density associated with the transition, while straight arrows represent the total magnetic (blue) and electric (red) dipoles. The black dashed arrows represent local magnetic moments for half-rotations of the orbital density.

Type A transitions are characterized by orbital density moving up or down the body of the helix, and involve MOs generally delocalized over several fused rings, which explains the parallel orientation of the magnetic dipole with the helical axis. At the same time, the MO centroids also move along the helix, so that the transition electric dipole moment is tilted with respect to the helical axis. The angle θ between the two dipoles is consistently $> 90^\circ$, corresponding to a negative $\tilde{S}_k$ value that increases the $[\alpha]_D$ magnitude (the two quantities are related through a minus sign, see Eq. 3.7). In this type of transition, the direction of the orbital density motion (up or down the helix body) is not important, as the angle between the electric and magnetic vectors is always obtuse, and the $\tilde{S}_k$ sign is consistently negative. Type B transitions can be interpreted as the combination of two separate rotations of the orbital density along the two halves of the helix.

Each half-rotation may be qualitatively described by a local magnetic vector that is tilted with respect to the main helical axis. In B_s transitions, where the terminal groups of the helicene are the same, the tilting angle has the same magnitude for both local magnetic vectors (i.e., ϕ and $\pi - \phi$), so that the components of these vectors parallel to the helical axis cancel out, and the resulting magnetic vector is perpendicular to the helical axis, see Figure 3.6b. In B_a transitions, where the terminal groups are different, the tilting angles of the local magnetic vectors are different in magnitude so that the cancellation of the components parallel to the helical axis is not perfect, and the resulting magnetic vector is also tilted, see Figure 3.6c. In either case, the electric vectors tend to be orthogonal to the main helical axis, or have similar orientations as in the Type A transitions. The result is that the overall angle θ between the electric and magnetic dipoles is always acute, thus the corresponding $\tilde{S}_k$ contributions decrease the $[\alpha]_D$ magnitude. However, given that the magnetic vectors are typically smaller in magnitude compared to those in the Type A transitions (due to the partial or complete cancellation of the components parallel to the helical axis in the vector sum of the local magnetic dipoles), the $\tilde{S}_k$ values for the Type B transitions are also smaller in magnitude than those for the Type A transitions. Hence, Type A transitions dominate the contribution to $[\alpha]_D$, and they determine its sign.

In the following sections, we compare the $[\alpha]_D$ values with the most important $\tilde{S}_k$ contributions in a single plot for each class of helicenes. Since $[\alpha]_D$ is an intensive property while the $\tilde{S}_k$ values are extensive quantities and they are opposite in sign, see Eq. 3.7, the plots report $-\tilde{S}_k/M$ rather than just $\tilde{S}_k$. Although $[\alpha]_D$ and $\tilde{S}_k$ have different units, the reported values are proportional through the C_ω factor in Eq. 3.7; in other words, the sum of all $-\tilde{S}_k/M$ terms would provide values coincident with the $[\alpha]_D$ values in the plots. The same figures also include the decomposition of each $\tilde{S}_k$ value in terms of the magnitude of the electric and magnetic dipoles as well as the cosine of the angle θ between the two vectors. The scope of these plots is to identify the main $\tilde{S}_k$ terms according to the transition type, and to understand which element of the dot product between the transition dipoles is responsible for the trend of $\tilde{S}_k$ (and ultimately $[\alpha]_D$) with N. The magnitude of the electric and magnetic dipoles as well as of $-\cos\theta$ are scaled by $M^{1/3}$ so that their product

equals the corresponding $-\tilde{S}_k/M$ value. Because all relevant transitions fall into one of the three categories discussed above (Type A, B_s, or B_a), we focus on the transitions with the largest $\tilde{S}_k$ value for the detailed discussion of the calculations on each set of molecules. For the figures that depict MO isodensity plots, we focus on a representative member of the set, since the MOs look similar across helicene lengths.

3.4.2.1 Unfunctionalized Helicenes

[N]helicenes are the most basic helicenes. The calculated $[\alpha]_D$ at this level of theory is in good agreement with experiment, as shown in Figure S10 in the SI.[109,120,121] Figure 3.7 reports plots of $[\alpha]_D$ for all unfunctionalized helicenes, including the main $\tilde{S}_k$ contributions and their sum, when more than one transition is important. A good qualitative agreement with the $[\alpha]_D$ trend is obtained with four transitions for [6]helicene (three Type A and one Type B_s), two transitions for [7]helicene (one Type A and one Type B_s), and five transitions for [8]helicene (three Type A and two Type B_s). An example of a Type A and a Type B_s transition is reported in Figures S11-S12 of the SI, which show the orbital isodensity and vector orientations.

As mentioned above, Type A transitions increase the magnitude of $[\alpha]_D$, while Type B_s transitions decrease it. For the larger helicenes (N = 7, 8), there is one very large Type A transition that dominates, which corresponds to a rotation of the orbital density over the entire body of the helix. The spatial extension of both the occupied and virtual MOs over the entire molecule is favored by the overlap of the terminal phenyl rings, see Figure S13 in the supporting information. Such overlap is not possible in the smaller helicene, so that the Type A $\tilde{S}_k$ contributions are all relatively small (2-4 a.u.); nevertheless, there is a considerable number of such transitions so that their collective contribution to $[\alpha]_D$ overcomes that of the large Type B_s $\tilde{S}_k$s. When we analyze the magnitude of the $\tilde{S}_k$ values in terms of the components of the vector dot product, we notice that the Type B_s transitions have similar values across the helicene set because the magnitude of the dipole vectors and the angle θ are not significantly influenced by the helical length. This is reasonable because increasing the number of phenyl ring units increases the molecular length only along the

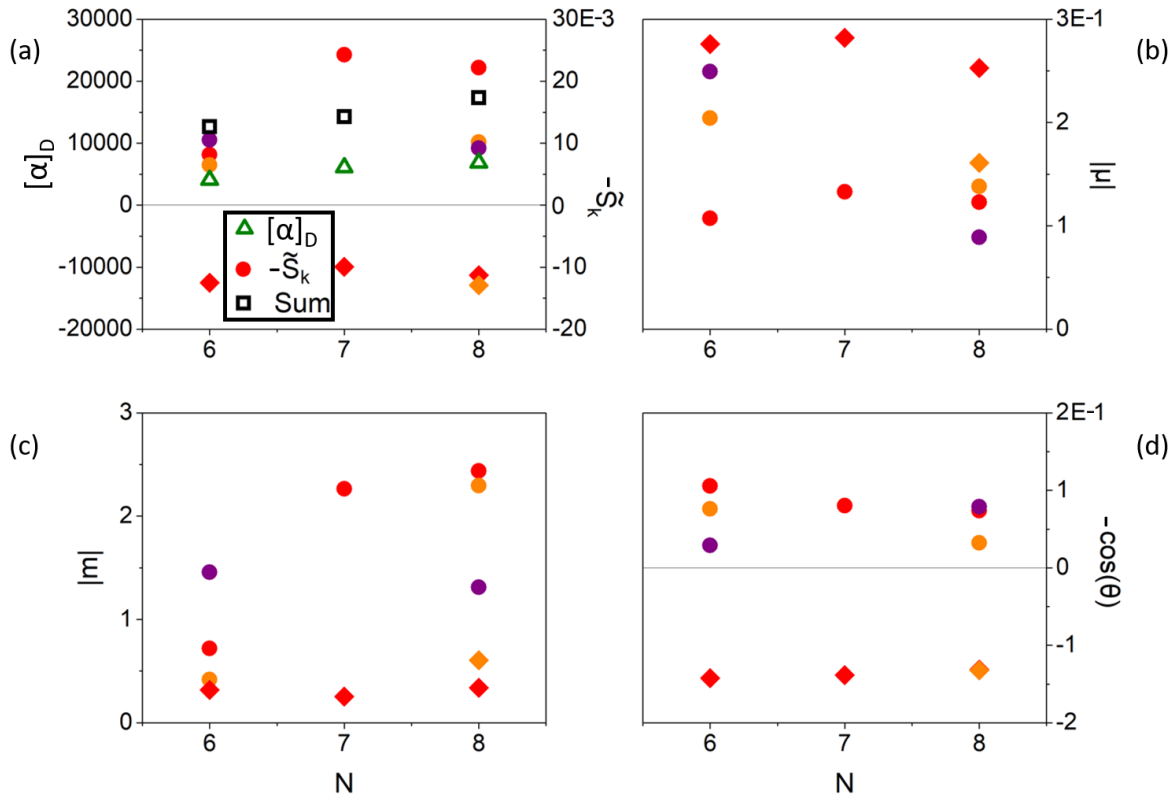


Figure 3.7: Detailed decomposition of $[\alpha]_D$ for [N]helicenes. a) $[\alpha]_D$ (deg [dm (g/mL)]$^{-1}$): green triangles; individual $-\tilde{S}_k/M$ values: circles (Type A) and diamonds (Type B$_s$), where colors distinguish transitions of the same type across subfigures; sum of $-\tilde{S}_k/M$ values: black squares. The other panels represent the electric dipole magnitude (b), magnetic dipole magnitude (c), and $-\cos\theta$ (d), where θ is the angle between the dipole vectors, for each transition, all scaled by $M^{1/3}$. The transitions presented in a) recover 103.5%, 78.0%, and 84.3% of the total $[\alpha]_D$ for N = 6–8, respectively.

main helical axis, but the vectors involved in the Type B_s transitions do not have any component along that direction, see Figure 3.6. On the other hand, the magnitude of Type A $\tilde{S}_k$ values is strongly correlated to the magnitude of the magnetic dipole, which increases considerably with the helicene length. At the same time, the electric dipole and the angle θ are not strongly influenced by the helical length because the latter does not affect the distance between the orbital centroids (at least for the range of N value considered in this work). Note that the electric dipoles for various transitions point in different directions depending on the relative position of the MO centroids, but the angle from the helical axis (and thus the angle θ with the corresponding magnetic vector) is rather constant.

It is well known that a parameter that determines the magnitude of $[\alpha]_D$ in helicenes is the helical pitch: the larger the pitch, the larger the $[\alpha]_D$ magnitude. We investigate this trend using [6]helicene, where we perform a constrained geometry optimization by fixing the distance between two C centers one full-turn apart at the equilibrium distance, and at ± 0.5 Å from equilibrium. The $\tilde{S}_k$ analysis shows that one leading Type A transition changes significantly (although the orbitals involved remain qualitatively of the same nature and shape), becoming more and more dominant with increasing pitch, see Figure 3.8. The $\tilde{S}_k$ decomposition clearly shows that its magnitude is mostly influenced by the angle θ between the electric and magnetic vectors, whose relative changes in magnitude largely compensate each other. More specifically, it is the direction of the electric dipole that determines the change in angle, because of the change in position of the MO centroids, while the direction of the magnetic dipole remains parallel to the helical axis. Thus, the change of $[\alpha]_D$ with the helical pitch is a geometric rather than an electronic effect.

3.4.2.2 Mono-Functionalized Helicenes

In this section and the next, we investigate the factors that determine the changes in $[\alpha]_D$ due to functionalization. As mentioned in Section 3.4.2, we use BTD as a π-electron withdrawing group, and two thiol units as π-electron donating groups.

The BTD group increases $[\alpha]_D$ compared to the unfunctionalized helicene analogues, as shown

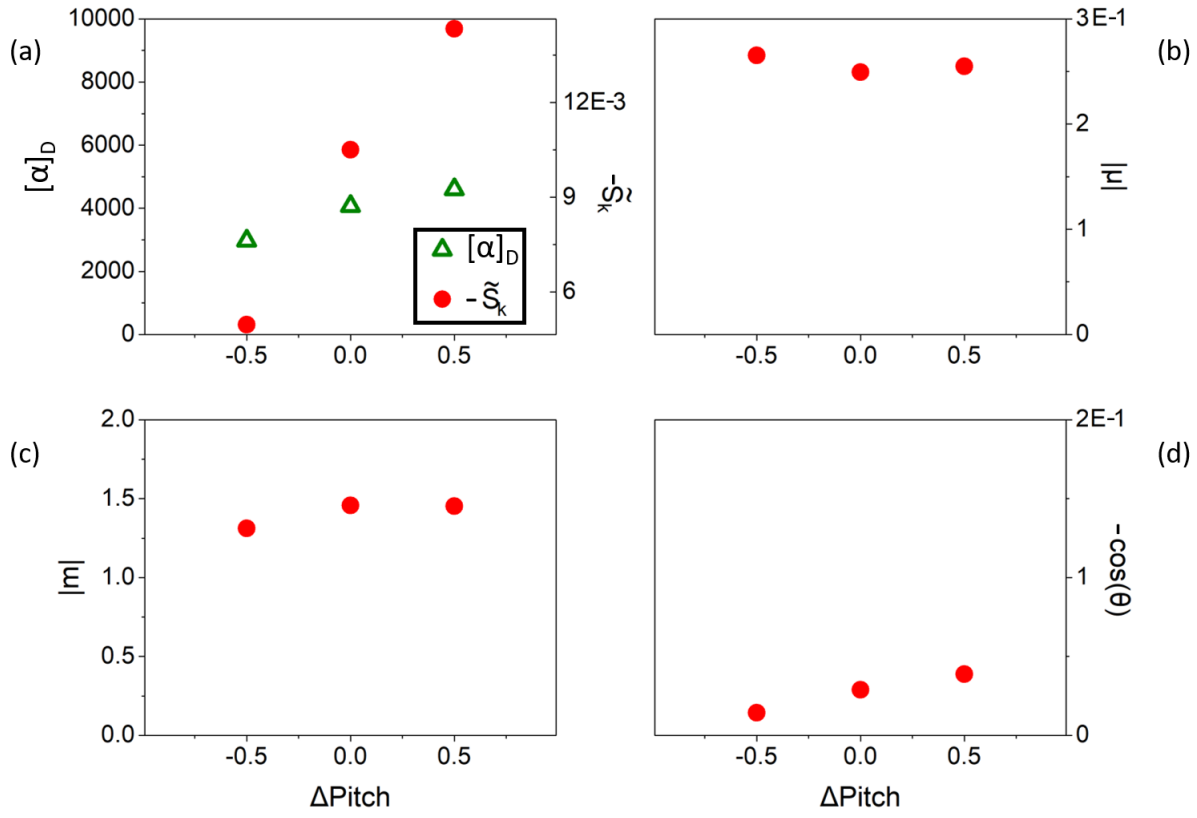


Figure 3.8: Detailed decomposition of $[\alpha]_D$ for [6]helicene with different pitch. a) $[\alpha]_D$ (deg [dm (g/mL)]$^{-1}$): green triangles; individual $-\tilde{S}_k/M$ values: circles (Type A). The other panels represent the electric dipole magnitude (b), magnetic dipole magnitude (c), and $-\cos\theta$ (d), where θ is the angle between the dipole vectors, for each transition, all scaled by $M^{1/3}$. The three transitions presented in a) recover 55.6%, 85.9%, and 103.9% of the total $[\alpha]_D$ for ΔPitch = -0.5, 0, +0.5 Å, respectively.

in Figure 3.5. A more detailed representation of the results for the BTD[N-2]helicenes is reported in Figure 3.9. This figure indicates that Type A transitions dominate the $\tilde{S}_k$ contributions to $[\alpha]_D$, and only one transition is necessary to reproduce the qualitative trend with N. The magnitude of the electric and magnetic dipoles is similar to that of [N]helicenes, with the former decreasing and the latter increasing with increasing N. At the same time, the angle between the vectors increases with N. This behavior can be explained by considering the orbitals involved in Type A transitions explicitly, and an example for BTD[6]helicene is shown in Figure 3.10. The presence of the electron withdrawing group tends to localize the virtual MO on the BTD unit, while the occupied MO is still delocalized across the helix, albeit to a lesser degree than in the unfunctionalized helicene analogue. Overall, the $\tilde{S}_k$ value for this transition is similar to that of [6]helicene. However, with increasing N, the more localized nature of the virtual MO simultaneously induces: 1) a longer rotation around the helix body compared to the corresponding [N]helicenes, thus slightly longer magnetic vectors by a factor of 1.2-1.45; 2) shorter electric dipole vectors, because the orbital centroids get closer to each other; and 3) larger angles θ, as the electric dipoles become more aligned with the helical axis while pointing away from the magnetic vector. At the same time, the asymmetry in the terminal groups of BTD[N-2]helicenes changes the Type B_s transitions into B_a transitions. The latter have $\tilde{S}_k$ values that are opposite in sign compared to the Type A transitions, and smaller in magnitude compared to both Type A and B_s transitions because the angle between the vectors is closer to 90°, see Figure 3.6. Taken together, strong Type A and weak type B_a contributions result in an overall increase in the $[\alpha]_D$ magnitude of BTD[N-2]helicene compared to [N]helicenes by a factor of 1.25-1.3.

Figure 3.5 shows that $[\alpha]_D$ for dithiol[N-1]helicenes is reduced by 5-20% compared to that of the [N]helicenes. Figure 3.11 indicates that a single Type A transition recoveres a significant portion of the $[\alpha]_D$ for each of the three helicenes, although reproducing the internal trend passing from N = 7 to 8 requires more than one transition. This is because the Type A $\tilde{S}_k$ values are overall smaller for the dithiol set than for the others, thus their relative importance compared to smaller $\tilde{S}_k$ contributions is decreased. Nevertheless, the examination of a single Type A transition is sufficient

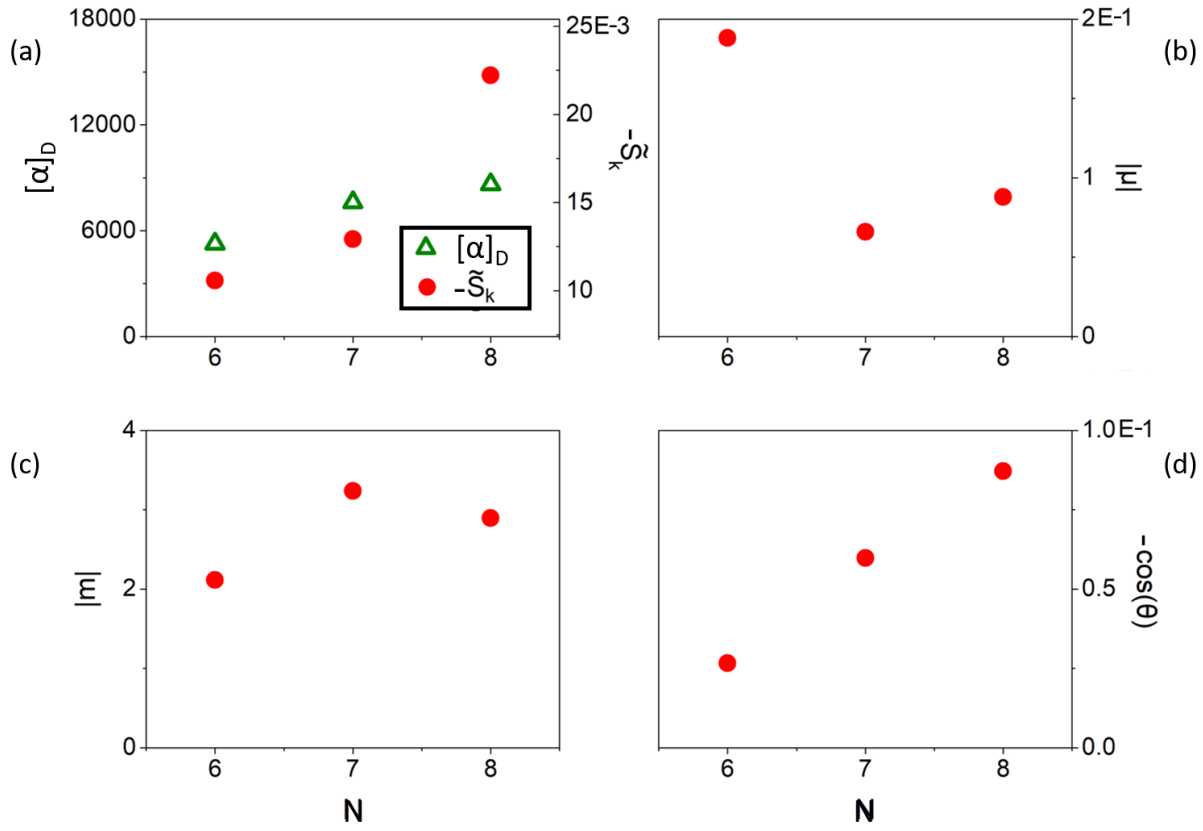
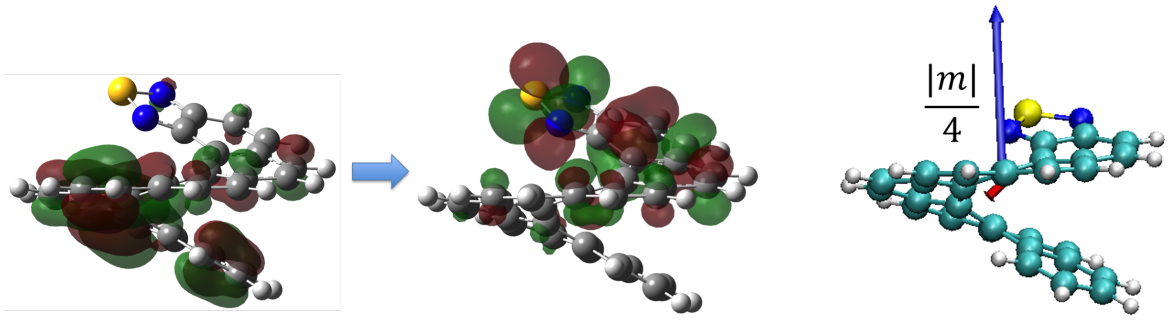


Figure 3.9: Detailed decomposition of $[\alpha]_D$ for BTD[N-2]helicenes. a) $[\alpha]_D$ (deg [dm (g/mL)]$^{-1}$): green triangles; individual $-\tilde{S}_k/M$ values: circles (Type A), where colors distinguish transitions of the same type across subfigures; sum of $\tilde{S}_k$ values: black squares. The other panels represent the electric dipole magnitude (b), magnetic dipole magnitude (c), and $-\cos\theta$ (d), where θ is the angle between the dipole vectors, for each transition, all scaled by $M^{1/3}$. The three transitions presented in a) recover 67.1%, 56.5%, and 85.8% of the total $[\alpha]_D$ for N = 6–8, respectively.



(a) MOs (b) Transition dipoles

Figure 3.10: MOs (a), and electric (red) and magnetic (blue) transition dipole vectors (b) for a Type A transition of BTD[6]helicene.

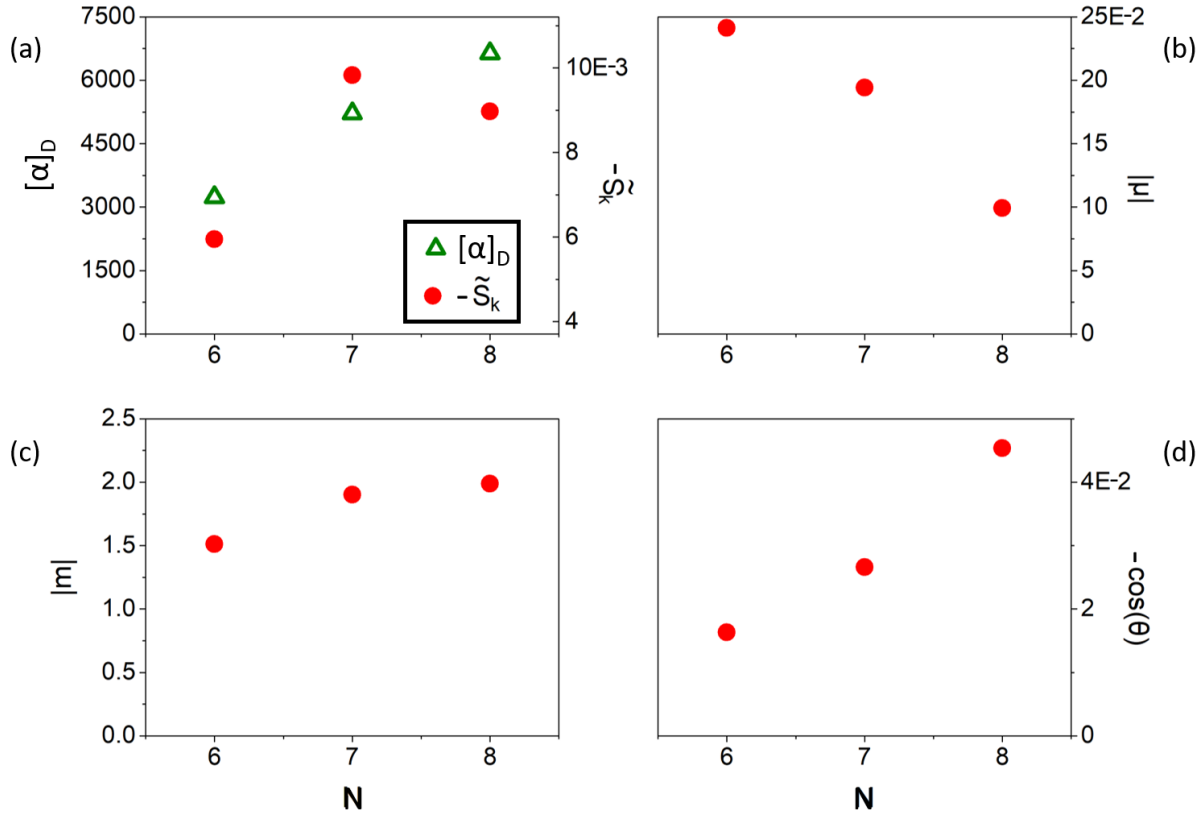
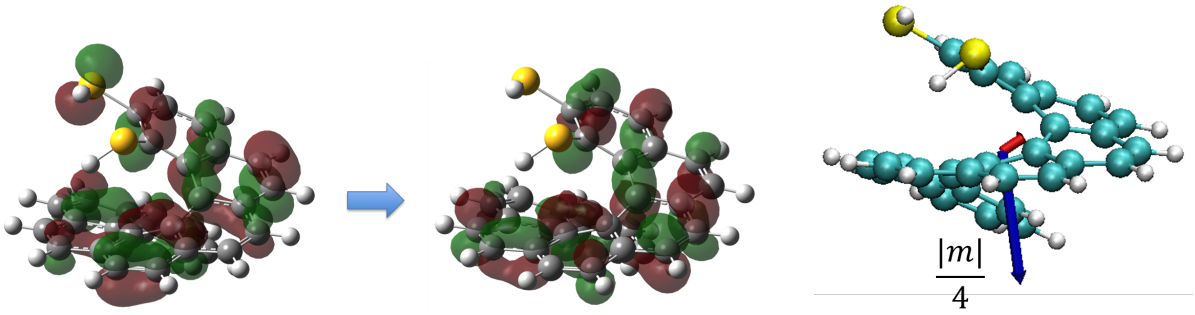


Figure 3.11: Detailed decomposition of $[\alpha]_D$ for dithiol[7]helicene. a) $[\alpha]_D$ (deg [dm (g/mL)]$^{-1}$): green triangles; individual $-\tilde{S}_k/M$ values: red circles (Type A). The other panels represent the electric dipole magnitude (b), magnetic dipole magnitude (c), and $-\cos\theta$ (d), where θ is the angle between the dipole vectors, for each transition, all scaled by $M^{1/3}$. The three transitions presented in a) recover 61.4%, 62.8%, and 45.0% of the total $[\alpha]_D$ for N = 6–8, respectively.

to explain the overall smaller values of $[\alpha]_D$ for the dithiol set compared to the unfunctionalized helicenes. The trends for the vector magnitudes and relative angles, also in Figure 3.11, are similar to those for BTD[N-2]helicenes, see Figure 3.9. The main difference is that the magnitude of the magnetic dipoles and the angle θ are smaller with the dithiol groups than with BTD. The reason for this change can be understood by considering the orbitals involved in a typical transition, shown in Figure 3.12. With dithiol, it is the occupied MO that is now more localized on the substituent group. However, since dithiol is a weaker electron donating group than BTD is an electron withdrawing group (compare the atomic charges of the bridging carbons of BTD, phenyl, and dithiol model systems in Figure S14 of the SI),[122,123] the localization is not as strong as in the BTD[N-2]helicenes. This leads to the smaller magnetic dipole and angle θ values, and ultimately $\tilde{S}_k$ and $[\alpha]_D$ values, even if the trends are consistent for both choices of substituents. Furthermore, as for BTD, Type B_a transitions also tend to decrease $[\alpha]_D$, and their relative importance is larger because the Type A $\tilde{S}_k$ values are smaller. Compared to [N]helicenes, the relative smaller magnitude of the magnetic dipoles of the dithiol[N-1]helicenes, and the compensating changes in the electric dipole magnitude and angle θ lead to overall smaller $[\alpha]_D$ values. However, the weak electron donating nature of the dithiol group also leads to a smaller deviation from the [N]helicenes $[\alpha]_D$ values compared to BTD[N-2]helicenes.



(a) MOs (b) Transition dipoles

Figure 3.12: MOs (a), and electric (red) and magnetic (blue) transition dipole vectors (b) for a Type A transition of dithiol[7]helicene.

3.4.2.3 Bis-Functionalized Helicenes

We now consider the functionalized helicenes with two substituent groups at both ends of the helix: bis-BTD[N-4]helicenes, bis-dithiol[N-2]helicenes, and mixed dithiol-BTD[N-3]helicenes. The corresponding $[\alpha]_D$ values are plotted in Figure 3.5. The specific rotation for the bis-BTD species with for N = 6, 7 is similar in magnitude to that of the corresponding mono-BTD species, but it is larger by 20% for the N = 8 compound. The detailed component separation for the $\tilde{S}_k$ values of the bis-BTD[N-4]helicenes is reported in Figure 3.13, where it is clear that one Type A transition dominates. Although for these transitions the magnetic dipole moment is parallel to the helical axis, they can be also described as two half-rotations of the orbital density from the central rings towards the terminal BTD groups, where the virtual MOs are localized, see Figure 3.14 for the N = 8 case. Contrary to regular Type B_s transitions, the overlap of the virtual MO region in between the BTD units can be described as a phase change for one of the half-rotations, which in turn changes the direction of the corresponding local magnetic dipole moment. Therefore, the components of the local magnetic dipoles that are parallel to the helical axis add up, while the component orthogonal to that axis cancel out, resulting in a total magnetic dipole parallel to the helical axis as in a typical Type A transition. The overlap of the BTD groups is particularly large for bis-BTD[4]helicenes (i.e., N = 8), which produces a strong pull for the half rotations, and a particularly large magnetic dipole moment and related $\tilde{S}_k$ value. On the other hand, the resulting magnetic dipole for the N = 6, 7 species is of the same magnitude of that for the mono-BTD species, while the changes in electric dipole magnitude and angle θ compensate each other. As a result, $[\alpha]_D$ for the mono and bis-BTD compounds with N = 6, 7 are similar in magnitude. Note that for bis-BTD[N-4]helicenes there are important contributions from Type B_s transitions, given that these compounds belong to the same point group as [N]helicenes, but these contributions are compensated by a number of smaller Type A $\tilde{S}_k$ terms (not shown).

Figure 3.15 reports the detailed data for the largest $\tilde{S}_k$ contributions for bis-dithiol[N-2]helicenes. Also in this case, one Type A transition dominates. These compounds are characterized by a $[\alpha]_D$ that is reduced further from the dithiol compounds by 50-70%, due to smaller Type A $\tilde{S}_k$ contribu-

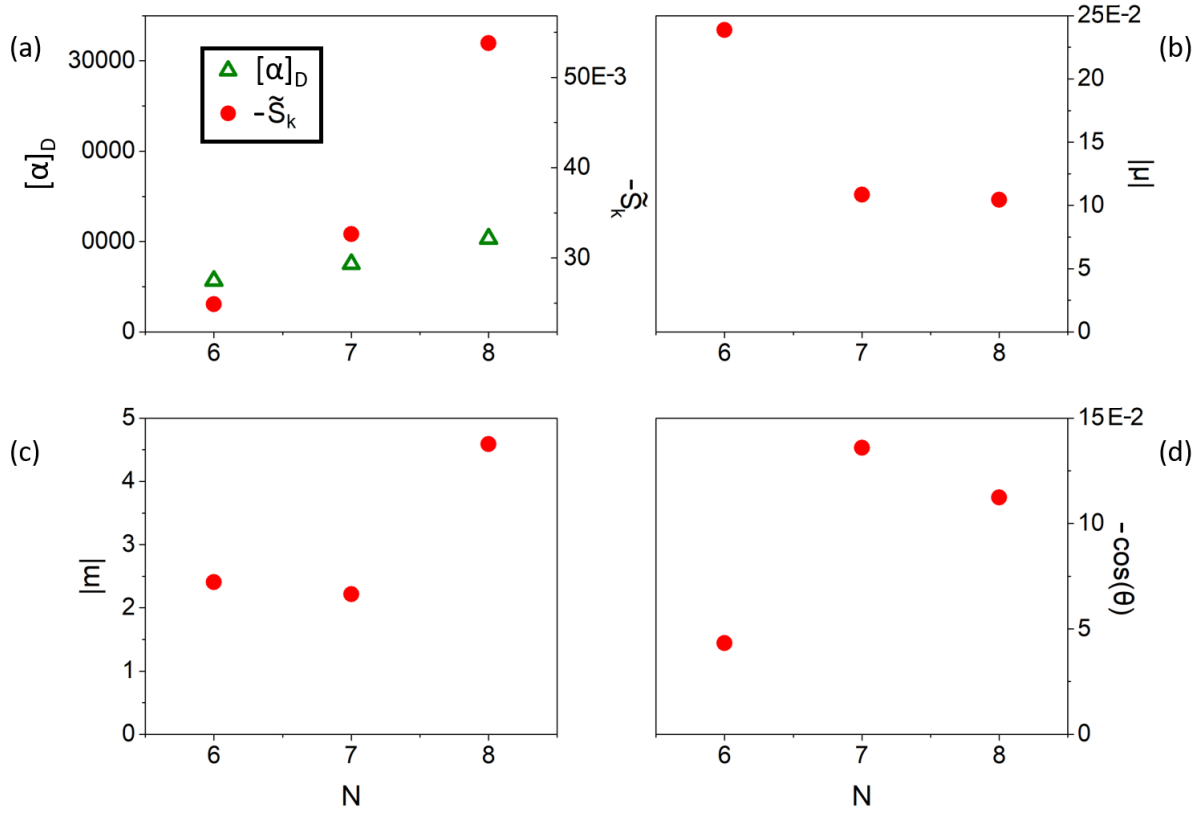
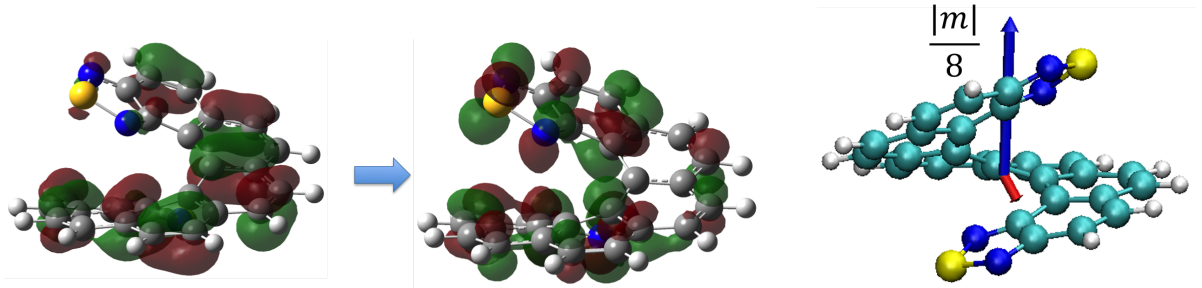


Figure 3.13: Detailed decomposition of $[\alpha]_D$ for bis-BTD[N-4]helicenes. a) $[\alpha]_D$ (deg [dm (g/mL)]$^{-1}$): green triangles; individual $-\tilde{S}_k/M$ values: red circles (Type A). The other panels represent the electric dipole magnitude (b), magnetic dipole magnitude (c), and $-\cos\theta$ (d), where θ is the angle between the dipole vectors, for each transition, all scaled by $M^{1/3}$. The three transitions presented in a) recover 146.6%, 144.7%, and 173.5% of the total $[\alpha]_D$ for N = 6–8, respectively.



(a) MOs (b) Transition dipoles

Figure 3.14: MOs (a), and electric (red) and magnetic (blue) transition dipole vectors (b) for a Type A transition of bis-BTD[4]helicene.

tions. These Type A transitions are similar in nature to those for the bis-BTD species, as shown in Figure 3.16 for a typical case, except that now it is the occupied MO that is mostly localized at the edges of the helix (i.e., on the dithiol group) while the virtual MO is more localized toward the central rings. However, the same discussion in terms of the sum of two half-rotations applies here as well to describe the contributions to the $\tilde{S}_k$ value. As for the mono-substituted helicenes, the dithiol group is a weaker directing group than BTD, so that the resulting magnetic dipole moments are smaller in magnitude. Although we limited the analysis to the largest Type A $\tilde{S}_k$ value, we note that a significant number of Type A and B_s transitions contribute to $[\alpha]_D$ with $\tilde{S}_k$ values that are similar in magnitude and opposite in sign, so that they mostly cancel out.

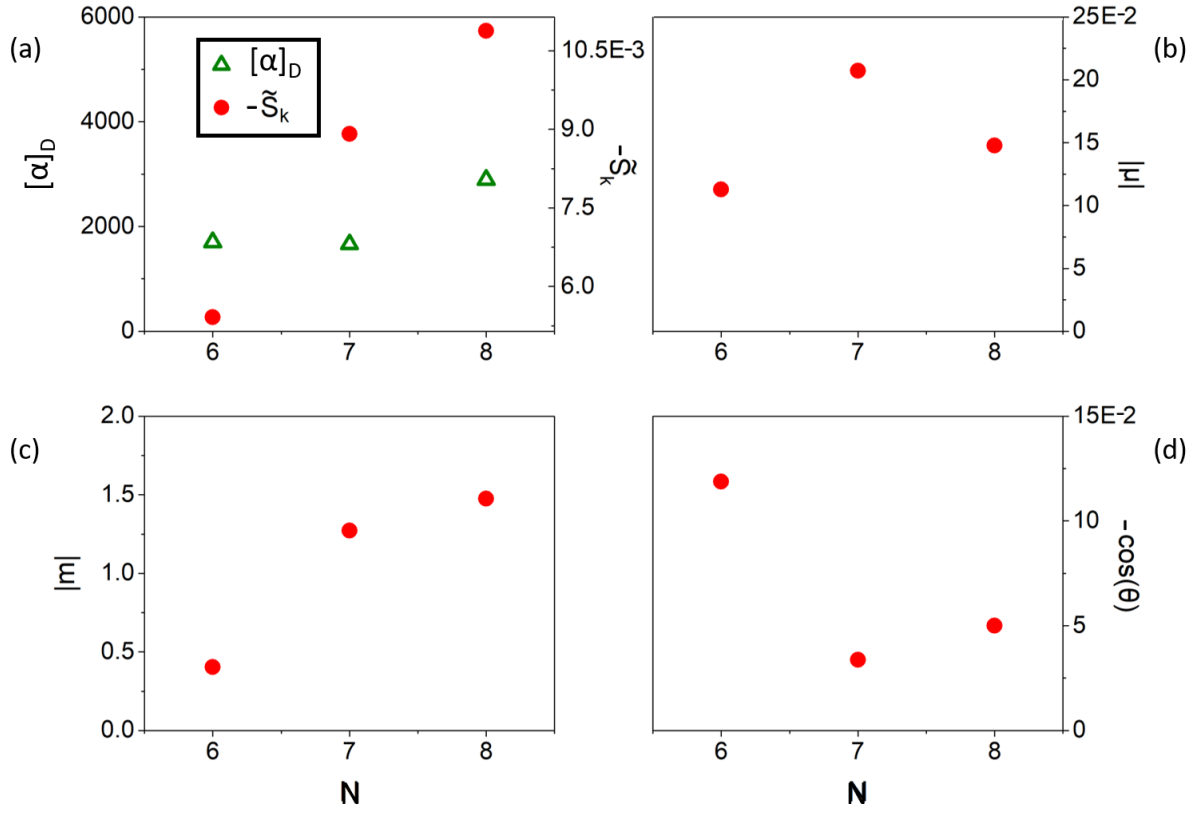
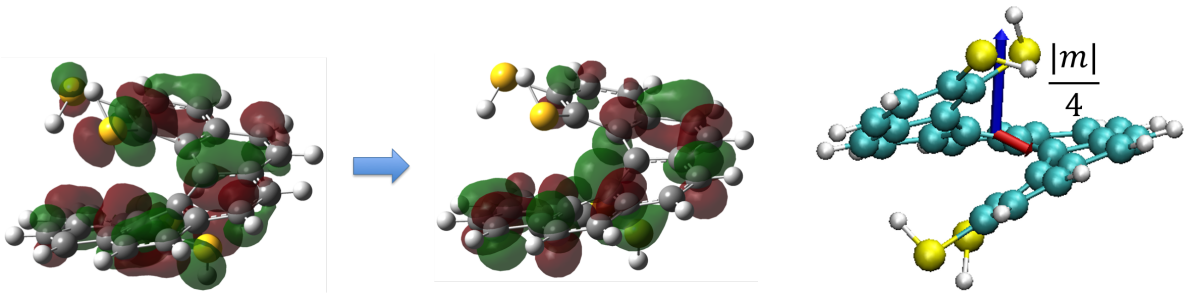


Figure 3.15: Detailed decomposition of $[\alpha]_D$ for bis-dithiol[N-2]helicenes. a) $[\alpha]_D$ (deg [dm (g/mL)]$^{-1}$): green triangles; individual $-\tilde{S}_k/M$ values: red circles (Type A). The other panels represent the electric dipole magnitude (b), magnetic dipole magnitude (c), and $-\cos\theta$ (d), where θ is the angle between the dipole vectors, for each transition, all scaled by $M^{1/3}$. The three transitions presented in a) recover 105.9%, 178.6%, and 125.4% of the total $[\alpha]_D$ for N = 6–8, respectively.

The $[\alpha]_D$ values for the mixed dithiol-BTD[N-3]helicenes are very close to those of [N]helicenes



(a) MOs (b) Transition dipoles

Figure 3.16: MOs (a), and electric (red) and magnetic (blue) transition dipole vectors (b) for a Type A transition of bis-dithiol[6]helicene.

(for N = 6, $[\alpha]_D$ is smaller by 13.5%, while for N = 7,8, $[\alpha]_D$ is smaller by only 1.5%), as reported in Figure 3.5. The trend of $[\alpha]_D$ with N for the dithiol-BTD compounds is reported in Figure 3.17 with the usual main contribution from a Type A transition. Similar to other cases, the trend in $\tilde{S}_k$ is determined by the length of the magnetic dipole, which increases with the length of the helix. However, the overall magnitude of the magnetic dipole is in general larger than the equivalent transitions in the unfunctionalized helicenes. The reason why the final $[\alpha]_D$ values are similar between these two sets of compounds is the angle θ, which is closer to 90° in the dithiol-BTD set. This is due to the asymmetry of the substituent groups, which affects the orbital density rotation around the helix body and results in a magnetic dipole that is slightly tilted away from the helical axis. This is an electronic effect rather than a geometrical one. It can be shown by replacing the dithiol groups in the dithiol-BTD[5]helicene with a regular fused phenyl unit, and reoptimizing the geometry of the new C and H centers while keeping the rest frozen (in other words, we constructed a BTD[6]helicene with the same pitch as the original dithiol-BTD[5]helicene). The magnetic dipoles in the original and modified structures are close in magnitude, but in the latter the dipole is less aligned with the helical axis, thus reducing the angle θ with the electric dipole, see Figure S15 in the SI.

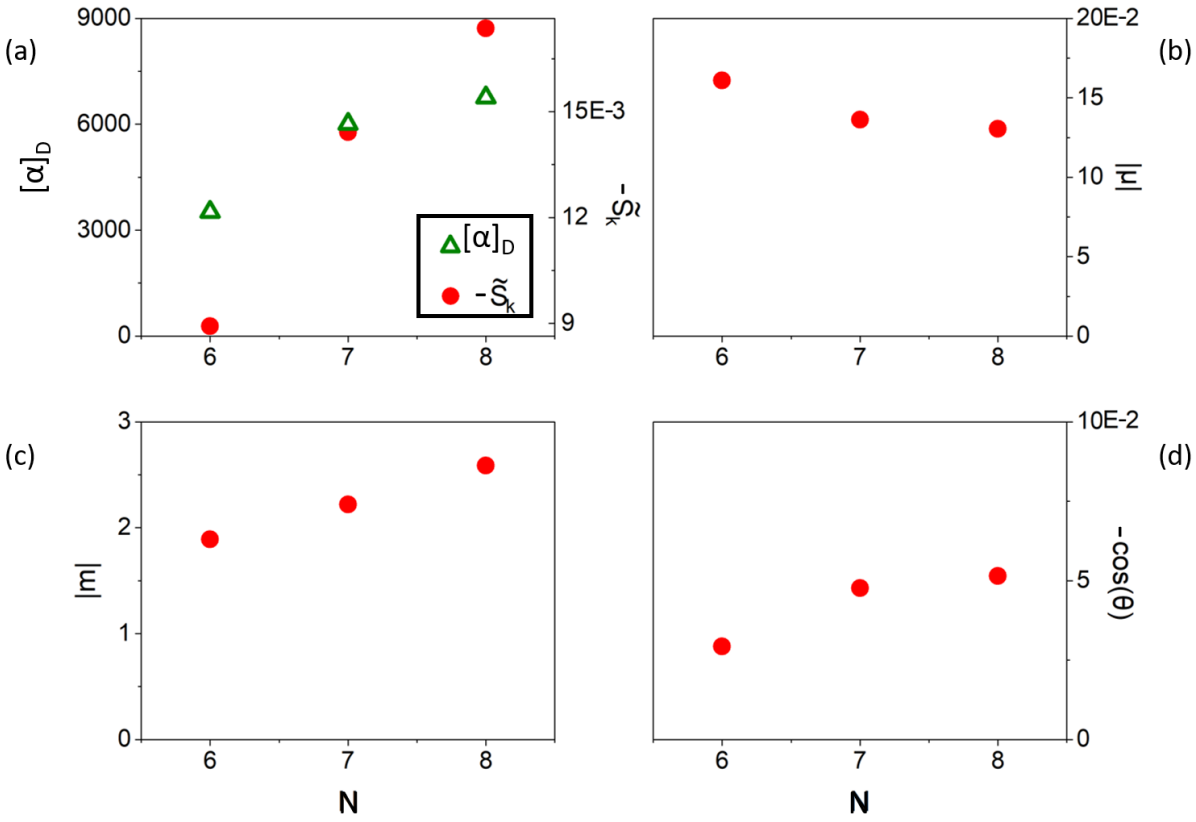


Figure 3.17: Detailed decomposition of $[\alpha]_D$ for dithiol-BTD[N-3]helicenes. a) $[\alpha]_D$ (deg [dm (g/mL)]$^{-1}$): green triangles; individual $-\tilde{S}_k/M$ values: red circles (Type A). The other panels represent the electric dipole magnitude (b), magnetic dipole magnitude (c), and $-\cos\theta$ (d), where θ is the angle between the dipole vectors, for each transition, all scaled by $M^{1/3}$. The three transitions presented in a) recover 84.6%, 80.0%, and 85.6% of the total $[\alpha]_D$ for N = 6–8, respectively.

3.5 Discussion and Conclusions

In this work, we present an analysis of $[\alpha]_\omega$ of helicenes in terms of $\tilde{S}_k$ contributions, which allow the decomposition of this property according to one-electron excitations between MOs. We show that such analysis can be performed with different choices of MO representation. For helicenes, canonical MOs are the best choice because delocalized π orbitals offer the most natural and compact decomposition of the specific rotation. However, localized MOs are likely better for molecules with small functional groups, where LMOs may provide a more chemically intuitive partitioning of $[\alpha]_D$.

We focus on [N]helicenes with $N = 6 - 8$ as well as analogous functionalized species, where electron-withdrawing and donating groups direct the movement of the electron density and therefore affect $[\alpha]_D$. The scope is to determine how the rotation changes with N and with the functionalization groups. We find that the major contributions to $[\alpha]_D$ come from three types of transitions, categorized depending on the orientation of the magnetic dipole term: parallel to the helical axis (Type A), orthogonal to the helical axis (Type B_s), and tilted (Type B_a), see Figure 3.6. Type A contributions are the strongest, and determine the sign of the rotation. Type B transitions are progressively weaker and have the opposite sign, thus reducing the magnitude of $[\alpha]_D$. In particular, Type B_s transitions are present in symmetric compounds (i.e, those with the same terminal groups and C_2 symmetry), while Type B_a transitions occur in asymmetric compounds (i.e., those with different terminal groups).

Type A transitions are mostly characterized by movement of the electron along the entire body of the helix, thus generating a large magnetic response along the helical axis. The magnitude of these transitions increases with N because the electron can move further, and their $\tilde{S}_k$ values are strongly dominated by the magnitude of the magnetic dipole. Terminal substituents affect this type of transition by a tendency to localize the occupied or virtual MOs on the group. For strong groups (BTD), this increases the extent of the motion along the body of the helix compared to the unfunctionalized helicene analogues, thus increasing the magnitude of the magnetic dipole and the corresponding $\tilde{S}_k$ value. In turn, this results in an overall increase of the specific rotation. For

weaker groups (dithiol), the partial localization reduces the extent of the motion along the body of the helix, thus resulting in smaller $[\alpha]_D$. Bis-functionalization induces Type A transitions with large $\tilde{S}_k$ values that may be described by simultaneous half-rotations of the electron from the terminal groups to the central rings, and vice-versa. In these half-rotations, the local magnetic dipoles sum constructively (thanks to a phase change in one of the MOs). The resulting $\tilde{S}_k$ values may be smaller, similar, or larger than those of the corresponding unfunctionalized helicenes depending on the strength of the terminal groups and the degree of overlap between these groups. Specifically, bis-dithiol substitution leads to smaller $[\alpha]_D$ than the [N]helicenes analogues, see Figure 3.15, bis-BTD substitution leads to similar values of $[\alpha]_D$ for N = 6, 7 and larger values of $[\alpha]_D$ for N = 8 (due to the overlap between the terminal BTD moieties), see Figure 3.13. Interestingly, a mixed substitution leads to $[\alpha]_D$ values similar to the [N]helicenes analogues, see Figure 3.17, because the magnetic dipoles of the large $\tilde{S}_k$ terms, although larger than in the unfunctionalized compounds, are tilted away from the helical axis due to the asymmetry of the terminal groups. The analysis of Type A transitions also allows us to show that the helical pitch effect on $[\alpha]_D$ is geometrical rather than electronic, as the magnitude of the electric and magnetic dipole vectors is mostly unchanged, but the angle between the two increases with the pitch, see Figure 3.8. Type B transitions can be described as two simultaneous half-rotations of the electron along the body of the helix, without phase change of the MOs. This results in complete (B_s) or partial (B_a) cancellation of the magnetic dipole component along the helical axis, which leaves the component perpendicular to the axis dominating. Therefore, these transitions correspond to smaller $\tilde{S}_k$ values than those of Type A. Additionally, the $\tilde{S}_k$ values of Type B transitions are not strongly dependent on N because the length of the helix only affects the contributions that are parallel to the helical axis.

It appears from this analysis that functionalization of the helicenes influences $[\alpha]_D$ in a complex manner, where the strength of the substituent groups and the length of the helix may lead to cooperative or competitive effects. At the same time, the $\tilde{S}_k$ analysis affords a relatively simple interpretation of these effects in terms of one-electron MO transitions. These studies may therefore lead to guidelines for the design of compounds with desired chiroptical responses.

3.6 Acknowledgement

The authors are grateful for support from the National Science Foundation under grant CHE-1650942.

3.7 Supporting Information

The Supporting Information includes the primary transition in the canonical MO (CMO) basis for (1S,4S)-norbornenone, plots of the cumulative contributions to $[\alpha]_D$ from $\tilde{S}_k$ for the N = 8 helicene derivatives with all MO bases, and the same plots for all helicenes with the CMO basis. Additionally, it contains a plot comparing experimental and calculated [N]helicene $[\alpha]_D$, examples of Type A and Type B_s transitions and MO overlap for [8]helicene, a comparison of the atomic charges of BTD, naphthalene, and dithiol models, and a comparison of the magnetic dipoles for dithiol-BTD[5]helicene and a modified BTD[6]helicene. Finally, it contains tables with the molecular weights of all helicenes, the numerical values of the data reported in Figures 3.7, 3.8, 3.9, 3.11, 3.13, 3.15, 3.17, as well as the Cartesian coordinates (Å) for the optimized structures.

Chapter 4

Comparison of Measured and Predicted Specific Optical Rotation in Gas and Solution Phases: A Test for the Polarizable Continuum Model of Solvation

(This work taken with the permission of Tal Aharon, Paul Lemler, Patrick Vaccaro, and Marco Caricato from *Chirality* **2018**, *30*, 383-395.[98] Supporting information is available online.)

4.1 Introduction

The study of chiroptical properties has been of great importance since the discovery of optical activity in crystals by Arago (1811) and Biot (1812), with continuing interest being due, in part, to the pervasive nature of chiral species in biochemical science and the pharmaceutical industry.[124] Since the specific optical rotation, $[\alpha]_{\lambda}^{T}$ (at wavelength λ and temperature T), of chiral molecules is tied closely to their absolute stereochemical configuration, which, in turn, often determines their pharmacological activity, ab initio studies of this fundamental quantity have become increasingly popular as they can provide direct access to a key structure-property correlation. Beginning with the work of Polavarapu in 1997 at the Hartree-Fock level of theory, quantum-chemical calculations of specific rotation have been extended to density functional theory (DFT) and to coupled cluster theory (CC).[5,7,9–14,16–20,125]

Although the aforementioned theoretical approaches successfully have aided in the characterization of chiral molecules and their interactions, open questions remain. One key issue is the ability to reproduce the effects of solvation on observed specific-rotation values. The work of the

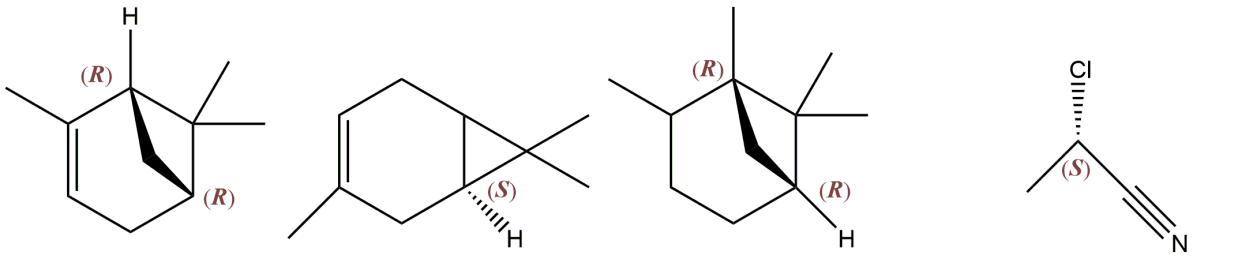
Vaccaro group on the measurement of dispersive chiroptical signatures in the vapor phase by means of cavity ring-down polarimetry (CRDP) has revealed unexpectedly large changes in the magnitude and even the sign of $[\alpha]_{\lambda}^{T}$ upon transferring chiral species from the vacuum to the condensed phase.[2,31,45,54,126–128] Despite extensive efforts, such large effects have eluded a robust theoretical characterization.

Several theoretical methods to approximate the effects of solvent upon $[\alpha]_{\lambda}^{T}$ have been reported. By combining continuum solvation models and molecular dynamics simulations, Mukhopadhyay *et al.* managed to obtain qualitative agreement with experimental data acquired for a simple chiral test molecule, methyloxirane. Their analyses showed the optical rotation in water was dominated by the solute contribution, while the dissymmetric first solvation shell formed in a benzene solvent contributed nearly as much to the specific rotation as the solute itself.[57,58] Kundrat *et al.* utilized molecular dynamics simulations to evaluate specific rotations for a series of amino acids. These authors employed explicit point charges to approximate solvent molecules in quantum-mechanical calculations created from classical-dynamics snapshots and compared emerging results to those obtained from a continuum-solvation model. They attributed the worse performance of the explicit solvation approach in reproducing the experimental $[\alpha]_{\lambda}^{T}$ values to shortcomings of DFT.[55,56] Crawford and co-workers applied CC calculations for the solute with frozen-density embedding (FDE) potentials for the solvent, but this approach was not sufficient to reproduce the solvation effect on specific rotation correctly, despite previous success with electronic absorption spectra.[54] Haghdani *et al.* used a microsolvation-plus-continuum scheme to calculate specific rotation in solution, but concluded further developments were needed to quantitatively predict solvation effects.[129] Other work has entailed use of polarizable force fields to describe the solvent, which showed promise for describing the change in sign of methyloxirane rotation observed in solution phases.[130] However, these approaches were unable to account consistently and generally for solvation effects on specific rotation, and they also required considerable computational effort as they all are based on multiple $[\alpha]_{\lambda}^{T}$ evaluations for tens or hundreds of snapshots derived from molecular dynamics simulations.

Another approach to describe solvation effects is based on continuum models, where the atomistic representation of the solvent is replaced by a continuous, polarizable medium. These methods provide considerable computational savings as they avoid the need for conformational sampling. One of the most widely used schemes is the polarizable continuum model (PCM) developed by Tomasi and co-workers.[131–134] Indeed, preliminary DFT-PCM studies have shown some success in reproducing solvent effects on $[\alpha]_{\lambda}^{T}$, suggesting that this may be an efficient scheme for predicting specific optical rotation in the condensed phase.[46–48] Recent efforts combining the linear response-CC (LR-CC) approach and PCM also have provided some encouraging results.[49] The present contribution systematically investigates the ability of PCM to reproduce solvent effects on the dispersive chiroptical properties exhibited by a series of conformationally rigid chiral molecules, and affords a thorough comparison of calculated and experimental values of specific optical rotation across multiple wavelengths and phases (viz., in the vapor phase and in various solvents possessing distinct polarities). In contrast to recent studies focusing on anomalously large rotatory powers imbued by the action of inherently dissymmetric chromophores, the ensuing analyses have addressed the more typical case of modest optical activities resulting from localized stereogenic elements that are perturbed asymmetrically by the surrounding chemical environment.[135,136] Zero-point vibrational corrections (ZPVCs) also are evaluated in the gas and solution phases, as they have been shown to be capable of significantly modifying the purely electronic response, and may be a dominating factor in the temperature dependence of the specific rotation.[30,37,41,42]

Measurements and calculations of specific rotation in both gas and solution phases are presented for a set of four molecules: (*R*)-α-pinene (**1**), (*S*)-3-carene (**2**), (*R*)-*cis*-pinane (**3**), and (*S*)-2-chloropropionitrile (**4**), the structures of which are shown in Figure 4.1. These compounds were chosen for their volatility and relative magnitude of rotation (which facilitate measurements in the vapor phase), for their solubility in organic solvents, and for their conformational rigidity (which avoids contributions from several isomers that would complicate comparisons between theory and experiment). The solvents selected for the present study are listed in Table 4.1, along with their static (ε_0) and optical (ε_∞) dielectric constants, which are the main parameters used in the

PCM treatment of solvation.



(*R*)-α-pinene (1) **(*S*)-3-carene** (2) **(*R*)-cis-pinane** (3) **(*S*)-2-chloropropionitrile** (4)

Figure 4.1: The set of chiral molecules targeted by the present study.

Table 4.1: Solvents used in this work, and their optical (ε_∞) and static (ε_0) dielectric constants.

Solvent	ε_∞	ε_0
Acetonitrile (ACN)	1.81	35.7
Methanol (MOH)	1.77	32.6
Chloroform (CHL)	2.09	4.7
Dibutyl Ether (DBE)	1.96	3.0
Benzene (BNZ)	2.25	2.3
Cyclohexane (CYH)	2.04	2.0

4.2 Materials and Methods

4.2.1 Theory and Computational Details

Calculations of $[\alpha]^T_\lambda$, expressed in deg dm^{-1} (g/mL)$^{-1}$, are reported, with the equilibrium values of these quantities (*vide infra*) being evaluated as:

$$\alpha_\lambda = \frac{1.92 \cdot 10^1 4\pi^2 N_A a_0^4}{M\lambda^2} Tr(\mathbf{G}') \tag{4.1}$$

where N_A is the Avogadro constant, a_0 is the Bohr radius in cm, λ is the wavelength of the incident light in nm, and M is the molecular weight of the chiral molecule in g/mol. Here $\mathbf{G}'$ is the mixed electric dipole-magnetic dipole polarizability tensor, or Rosenfeld tensor, which is computed using

standard linear-response techniques.[91,137–139] Zero-point vibrational corrections (ZPVCs) are evaluated by using a Taylor-series expansion of the specific rotation about the equilibrium geometry of the molecule:

$$[\alpha]_\lambda^T \approx [\alpha]_\lambda^{eq} + \frac{1}{2}\sum_{i=1}^{3N-6} \frac{\partial^2 [\alpha]_\lambda^{eq}}{\partial Q_i^2} \langle \Delta x_i \rangle_T^2 \tag{4.2}$$

where $[\alpha]_\lambda^{eq}$ is the specific rotation at the equilibrium geometry as defined by Eq. 1. This expansion neglects the first derivative of specific rotation with respect to the normal modes;[42] however, computation of this missing term also requires the evaluation of cubic force constants, which was deemed too demanding in the context of the present work. Therefore, the ZPVCs in Eq. 2 are limited to quadratic terms as this should be sufficient for a semi-quantitative discussion of the results. The sum in Eq. 4.2 runs over the harmonic vibrational degrees of freedom of the molecule,

with the square of the average thermal displacement along the i-th normal mode,$\langle \Delta x_i \rangle_T$, being given by:

$$\langle \Delta x_i \rangle_T^2 = \Big(\frac{16.8576}{\tilde{\nu}_i}\Big) \coth \frac{0.719384 \tilde{\nu}_i}{T} \tag{4.3}$$

where $\tilde{\nu}_i$ is the fundamental wavenumber (in cm-1) for the i-th vibrational displacement.[30] Only corrections at 0 K are considered, as this again is enough for a semi-quantitative discussion. The second derivative of the specific rotation in Eq. 4.2 is approximated as:

$$\frac{\partial^2 [\alpha]_\lambda^{eq}}{\partial Q_i^2} \approx \frac{[\alpha]_\lambda^{+\delta_i} - [\alpha]_\lambda^{eq} + [\alpha]_\lambda^{-\delta_i}}{\delta_i^2} \tag{4.4}$$

Here the mass-weighted displacement δ_i is obtained by normal-mode scaling:

$$\delta_i = \frac{Q_i s}{\sqrt{\tilde{\nu}_i}} \tag{4.5}$$

where s = 0.1 follows from the recommendation of Crawford and coworkers, and the square root

of the vibrational frequency in the denominator (in analogy to the average displacement in Eq. 4.3) is introduced to produce larger displacements for soft modes.[35]

All calculations were performed with a development version of the GAUSSIAN suite of programs.[114] Optimized geometries obtained by applying the B3LYP/aug-cc-pVTZ model chemistry in each respective medium (*i.e.*, gas phase and solution phase) were used for specific-rotation calculations with all other levels of theory to avoid geometrical effects. The $[\alpha]_\lambda^T$ calculations were performed with the aug-cc-pVDZ basis set, which offers a good compromise between computational accuracy and cost, at the following levels of theory: B3LYP, CAM-B3LYP, and coupled cluster singles and doubles (CCSD).[115,117–119,140–144] The DFT methods exploited the length gauge with gauge-including atomic orbitals (GIAOs) to remove origin dependence.[19,94,95] Since GIAOs cannot be used with standard CCSD, calculations are reported for the modified velocity gauge (MVG, origin independent) and the length gauge (LG, origin dependent) because the latter has shown better convergence with basis-set size.[17] All ZPVC calculations were performed with B3LYP, and the results were added to the equilibrium $[\alpha]_\lambda^T$ values computed at other levels of theory.

Calculations in solution used the symmetric integral equation formalism version of the polarizable continuum model (IEFPCM, but denoted as "PCM" in the following text).[145] In this model, the solvent is represented by a continuous and polarizable medium where a cavity of proper molecular shape hosts the solute. The solute-solvent electrostatic interaction is introduced through an effective term in the Hamiltonian, which depends on the macroscopic static dielectric constant, ε_0. PCM also can be used to introduce solvent effects on frequency-dependent molecular properties such as the specific rotation.[131,134] These calculations are performed in the non-equilibrium regime, which assumes the solvent molecules and nuclei are too slow to respond to changes in the solute electron density and are kept fixed (*i.e.*, the inertial response), while the solvent electrons respond instantaneously (*i.e.*, the dynamic response). The latter enters the solute linear-response equations used to evaluate the Rosenfeld tensor, and the PCM term now depends on the optical dielectric constant of the solvent, ε_∞ (*i.e.*, the square of the refractive index). The cavity is built as a series of interlocking spheres centered on the solute nuclei and the radii are parameterized. Based

on preliminary tests, cavities were built with Bondi radii and a scaling factor of 1.2 (*viz.*, R_C = 1.7 Å, R_H = 1.2 Å, R_N = 1.55 Å, and R_{Cl} = 1.75 Å). Added spheres have been used for the specific rotation calculations to avoid unphysical solvent pockets. ZPVC calculations were performed by maintaining the cavity fixed during numerical differentiation (*cf.* Eq. 4), as initial tests showed numerical instabilities when the cavity was moved during the geometry displacements.

4.2.2 Experimental Details

(All measurements in this work were performed by the Vaccaro group at Yale University)

Experimental values of specific optical rotation, $[\alpha]_\lambda^T$ in deg dm^{-1} (g/mL)$^{-1}$, at incident wavelength λ and near-ambient temperature T were obtained, in part, from published CRDP studies, which gave requisite intrinsic (vapor-phase) chiroptical properties for (*R*)-α-pinene (**1**),[126] (*S*)-3-carene(**2**),[31] (*R*)-*cis*-pinane (**3**),[126] and (*S*)-2-chloropropionitrile (**4**)[128] at λ = 355nm and λ = 633 nm. Complementary solvated quantities for **4** were acquired from the same source, while new solution-phase measurements were performed for other targeted species.[128]

Samples of **1**, **2**, and **3** were procured from a commercial source (Sigma-Aldrich) and used without further processing at their specified purities of 99%, 99%, and 98.5%. While **1** was stated by the manufacturer to have a percentage enantiomeric excess (%ee) of 97%, neat optical rotation measurements for **2** at the sodium D-Line (589.3 nm) suggested >97% ee when compared to previously published metrics and chiral GC-MS analyses of **3** (based on a Sigma-Aldrich B-DM column; 40 m length x 0.25 mm diameter) indicated >95% ee.[31] As such, explicit corrections of measured $[\alpha]_\lambda^T$ parameters for enantiomeric and chemical purity were deemed to be unnecessary.

Solution-phase studies of dispersive optical activity were performed in a temperature-regulated (25±1 °C) quartz sample cell (10.000 ± 0.005 cm length) by utilizing a commercial polarimeter (*Perkin-Elmer* 241; ± 0.002° angular accuracy) that operated at discrete visible/ultraviolet excitation wavelengths filtered from NaI and HgI atomic-emission lamps (*viz.*, 365.02 nm, 436.83 nm, 546.07 nm, 578.39 nm, and 589.30 nm). Each measurement was repeated at least twice with a substantial integration time (20 s) being used to ensure reproducibility, leading to reported quan-

tities that represent the average of acquired experimental results. Solvents selected for the present study, acetonitrile (ACN), methanol (MOH), chloroform (CHL), di-*n*-butyl ether (DBE), benzene (BNZ) and cyclohexane (CYH), were of spectrometric grade, and solute concentrations were kept as low as possible (typically $\leq 10^{-3}$ g/mL) to minimize aggregation (solute-solute) effects while still retaining high polarimetric precision.

Since intrinsic rotatory powers were restricted to excitation wavelengths that did not overlap with those available for solution-phase measurements, solvated results were extrapolated to their vapor-phase counterparts by performing nonlinear least-squares regressions based on the following functional form:

$$\begin{aligned}\left[\alpha\right]_T^{\lambda} &= A\frac{2\lambda_{eg}^2\left[\lambda^2-\lambda_{eg}^2\frac{\lambda\lambda_{eg}}{\gamma_{eg}}\right]}{\left[(\lambda-\lambda_{eg})^2+(\frac{\lambda\lambda_{eg}}{\gamma_{eg}})^2\right]\left[(\lambda+\lambda_{eg})^2+(\frac{\lambda\lambda_{eg}}{\gamma_{eg}})^2\right]} \\ &= A\Big(\frac{2\gamma_{eg}\lambda_{eg}}{(\lambda_{eg}-\gamma_{eg})^2}\Big)\frac{2\gamma_{eg}^2\lambda_{eg}^2\lambda^2-2\gamma_{eg}^2\lambda_{eg}^2\lambda^2(\lambda+\lambda_{eg})}{\lambda^4\lambda_{eg}^4+\gamma_{eg}^4(\lambda^2-\lambda_{eg}^2)+2\gamma_{eg}^2\lambda_{eg}^2\lambda^2(\lambda+\lambda_{eg}}\end{aligned} \tag{4.6}$$

where A is an overall amplitude factor while λ_{eg} and γ_{eg} denote the resonant wavelength and spectral dephasing linewidth for an isolated electronic transition between excited state $|e\rangle$ and ground state $|g\rangle$. For an $|e\rangle \leftrightarrow |g\rangle$ resonance characterized by dephasing rate $\Gamma_{eg} = \frac{1}{\tau}$ (in s$^-$1), where τ (in s) denotes the corresponding dephasing time, the quantity λ_{eg} (in m) is related formally through the speed of light in a vacuum, c, as $\lambda_{eg} = 2\pi c/\Gamma_{eg} = 2\pi c\tau$. Deduced from a full perturbative expansion for attendant matter-field interactions, this expression correctly predicts the vanishing magnitude of specific rotation in the asymptotic limit of long wavelengths ($\lambda \to \infty$).[2] In view of the three adjustable parameters (A, λ_{eg} and γ_{eg}) in Eq. (6), attempts to utilize this model directly for the least-squares interpolation of vapor-phase measurements were hindered by the availability of only two data points (at 355 nm and 633 nm). Under such circumstances, the value of A was constrained to equal that obtained from analyses of the closest solution-phase mimic for isolated-molecule behavior (often acetonitrile).

4.3 Results and Discussion

(All measurements in this work were performed by the Vaccaro group at Yale University)

The optical-activity results are presented as a series of correlation plots between experiment and theory. Two quantities of particular interest are considered: the equilibrium value of the specific rotation $[\alpha]_{\lambda}^{T}$, and the attendant solvation shift, $\Delta[\alpha]_{\lambda}^{T}$, defined by:

$$\Delta[\alpha]_{\lambda}^{T} = [\alpha]_{\lambda,solv}^{T} - [\alpha]_{\lambda,gas}^{T} \tag{4.7}$$

ZPVCs only are considered in a second stage because the effects on $\Delta[\alpha]_{\lambda}^{T}$ are large and would complicate the discussion. The aforementioned $[\alpha]_{\lambda}^{T}$ and $\Delta[\alpha]_{\lambda}^{T}$ plots are presented separately for each molecule and level of theory, with the $[\alpha]_{\lambda}^{T}$ results (gas and solution phase) for all wavelengths reported on the same graph. In this way, the performance across various media can be assessed readily via visual inspection. All numerical values are listed in Tables S26-S65 in the supporting information (SI). Four key wavelengths are examined in detail by the present study: 355 nm, 436.83 nm, 589.3 nm, and 633 nm, where 355 nm and 633 nm are the incident wavelengths available for gas-phase CRDP measurements, while 436.83 nm and 589.3 nm are two of the emission lines used for solution-phase measurements. Phase-dependent values of specific rotation at wavelengths for which measurements are not available are obtained from the least-squares regression procedure described in the Experimenal Details section (*cf.* Eq. 4.6). Experimental findings are reported along the horizontal axis, while the corresponding calculated quantities are reported along the vertical axis. In addition, two linear fits of each dataset have been performed, one for the gas phase and one for the solution phase (the latter considering all solvents concurrently).

Figure 4.2 displays optical-activity results for compound **1**. Experimentally, $[\alpha]_{\lambda}^{T}$ decreases when going from a rarefied gaseous medium to the condensed (solution) phase, except in methanol, for which the $[\alpha]_{\lambda}^{T}$ increases. The slope of the CAM-B3LYP linear fit shows a nearly 1:1 correspondence between theory and experiment for the gas phase, with experiment being over-predicted for the solution phase. B3LYP and CCSD-LG underestimate both gas and solution phase measure-

ments. CCSD-MVG calculations predict $[\alpha]_\lambda^T$ values that are too low in the gas phase, but reproduce solution-phase measurements and both DFT methods are contradictory to what is found in experiments. For instance, the calculated solvent effects for acetonitrile and chloroform are large while the experimental shifts are small (measured 355 nm values of 188.2, 186.6, and 188.8 deg dm^{-1} (g/mL)$^{-1}$ in gas, ACN, and CHL). In addition, experiments performed in benzene, cyclohexane, dibutyl ether, and methanol show a significant change, amounting to between 10-20% of the gas phase value (measured 355nm values of 188.2, 146.3, 152.8, 166.0, and 197.1 deg dm^{-1} (g/mL)$^{-1}$ in gas, BNZ, CYH, DBE, and MOH), while the best predicted shifts for the same solvents, obtained with PCM and CCSD-MVG, range between 5-10% (CCSD-MVG 355 nm values of 157.9, 138.7, 152.8, 143.0, and 168.8 deg dm^{-1} (g/mL)$^{-1}$ in gas, BNZ, CYH, DBE, and MOH). The experimental $[\alpha]_\lambda^T$ parameters decrease in magnitude from acetonitrile to benzene in order of decreasing ε_0 and increasing ε_∞ (*cf.* Table 1). In all calculations, this ranking is maintained for acetonitrile, cyclohexane, and benzene but not for dibutyl ether and methanol. While the correlation between experimental and calculated rotatory powers increases linearly with decreasing wavelength for CCSD and CAM-B3LYP, the B3LYP data points deviate markedly from linearity.

For compound **2**, Figure 3 reveals large differences between gas-phase and solution-phase results. The correlation plots show very good agreement for gas-phase rotatory powers across all methods, although experimental values consistently are underestimated. Similarly, the comparison between theory and experiment for solvated response is fairly linear across all methods; however, predicted values generally overestimate observed behavior. Interestingly, the experimental findings suggest a large shift in $[\alpha]_\lambda^T$ between isolated and solvated conditions, indicating the action of a general "solvation" shift that is nearly independent of the choice of solvent. The relative change in magnitude between gas-phase chiroptical response and that of the closest-lying solvent, dibutyl ether, is five times greater than the largest difference between two solvents (measured 355nm values of 153.8, 85.5, and 72.1 deg deg dm^{-1} (g/mL)$^{-1}$ in gas, DBE, and CHL). The sign of this shift is not in keeping with calculations, which suggest $[\alpha]_\lambda^T$ to be larger in solution than in the gas phase (*cf.* ensuing discussion and $[\alpha]_\lambda^T$ plots). Experimental solvent trends also are not reproduced,

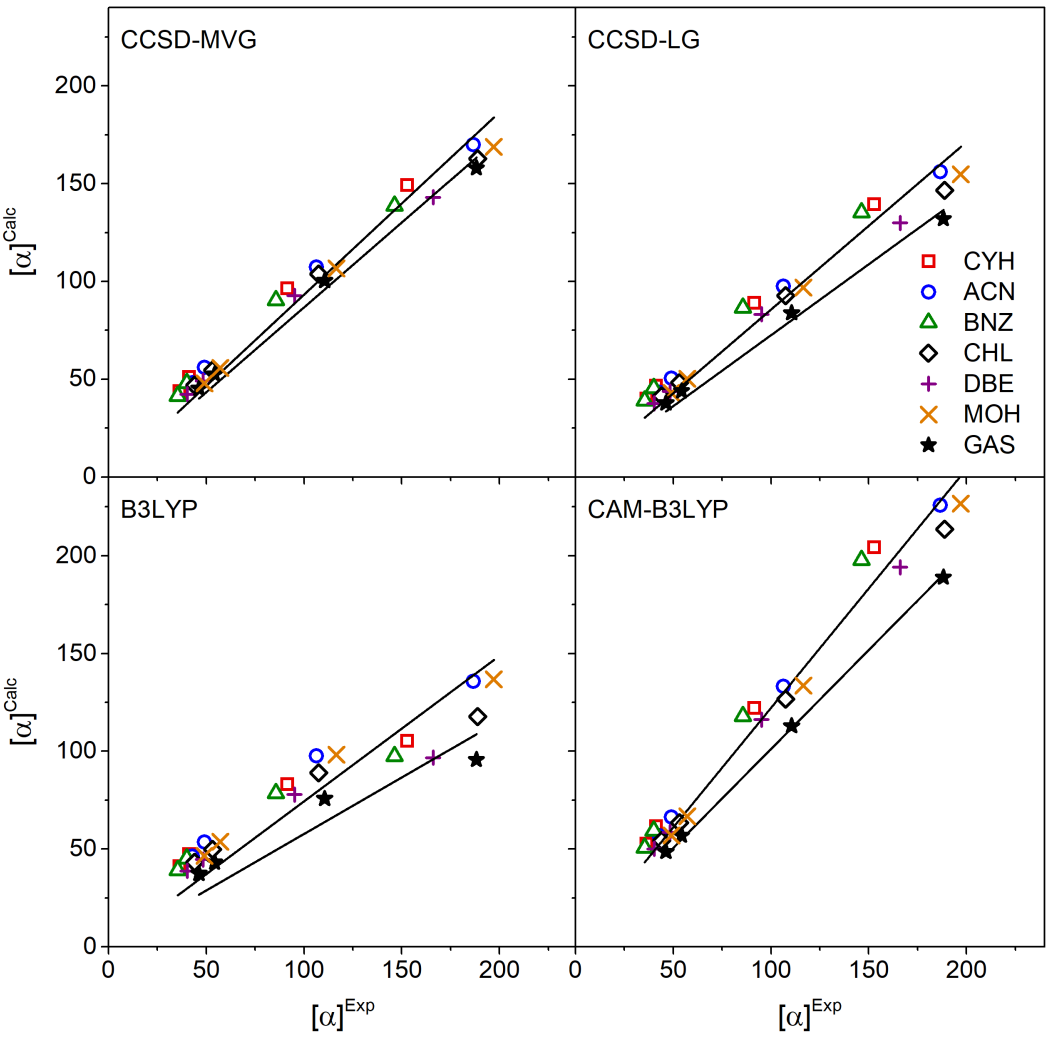


Figure 4.2: Gas and solution phase plots of $[\alpha]^T_\lambda$ (in deg dm^{-1} (g/mL)$^{-1}$) for (*R*)-α-pinene (**1**). The solvent acronyms are listed in Table 1. Experimental results are provided by the Vaccaro Group.

as the medium with the largest observed rotatory powers, dibutyl ether, has the smallest predicted $[\alpha]_{\lambda}^{T}$ values (across all methods), while the opposite behavior is obtained for chloroform.

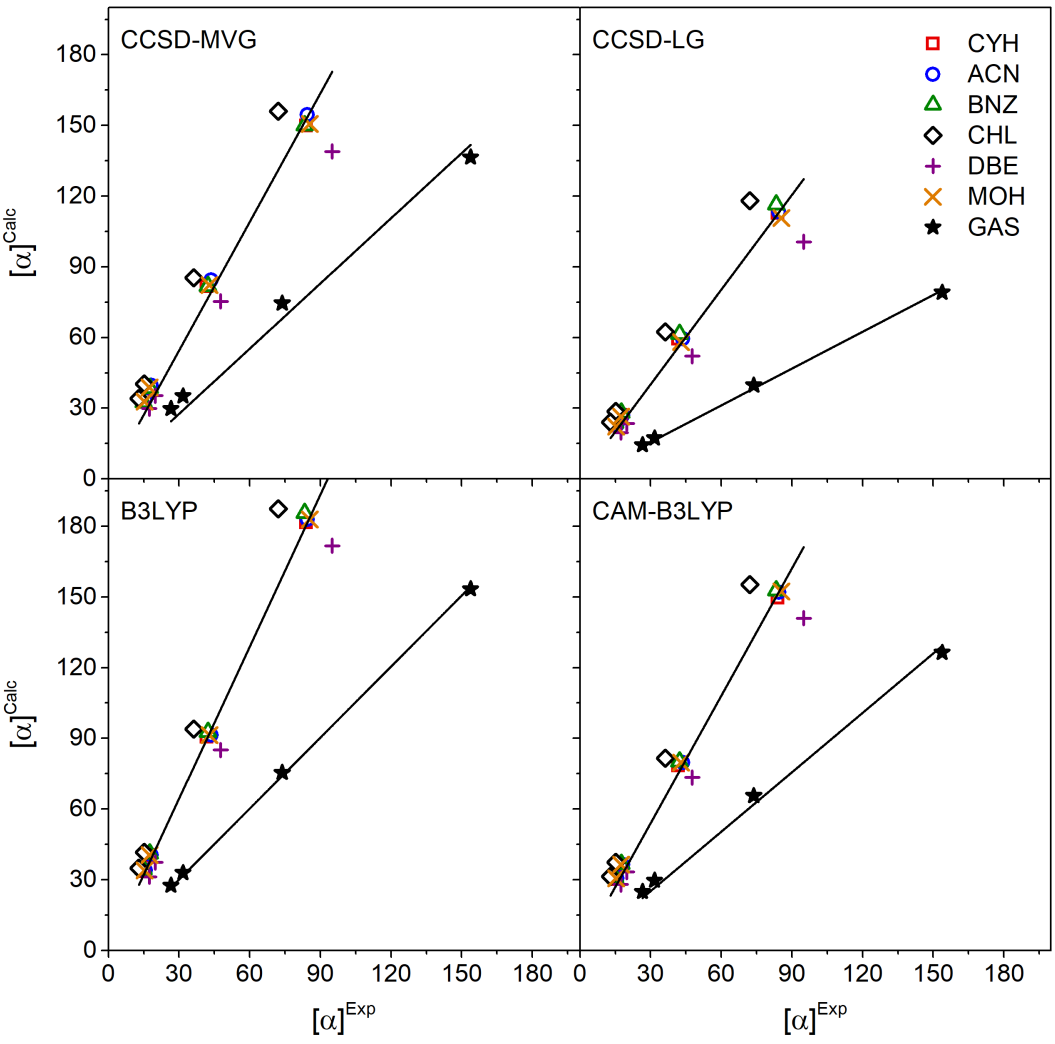


Figure 4.3: Gas and solution phase plots of $[\alpha]_{\lambda}^{T}$ (in deg dm^{-1} (g/mL)$^{-1}$) for (*S*)-3-carene (**2**). The solvent acronyms are listed in Table 1. Experimental results are provided by the Vaccaro Group.

The calculated and experimental $[\alpha]_{\lambda}^{T}$ parameters for compound **3** are highlighted in Figure 4. As with the previous systems, the correlation between theory and experiment under rarified conditions is quite good, with DFT showing particularly strong results indicated by a linear slope

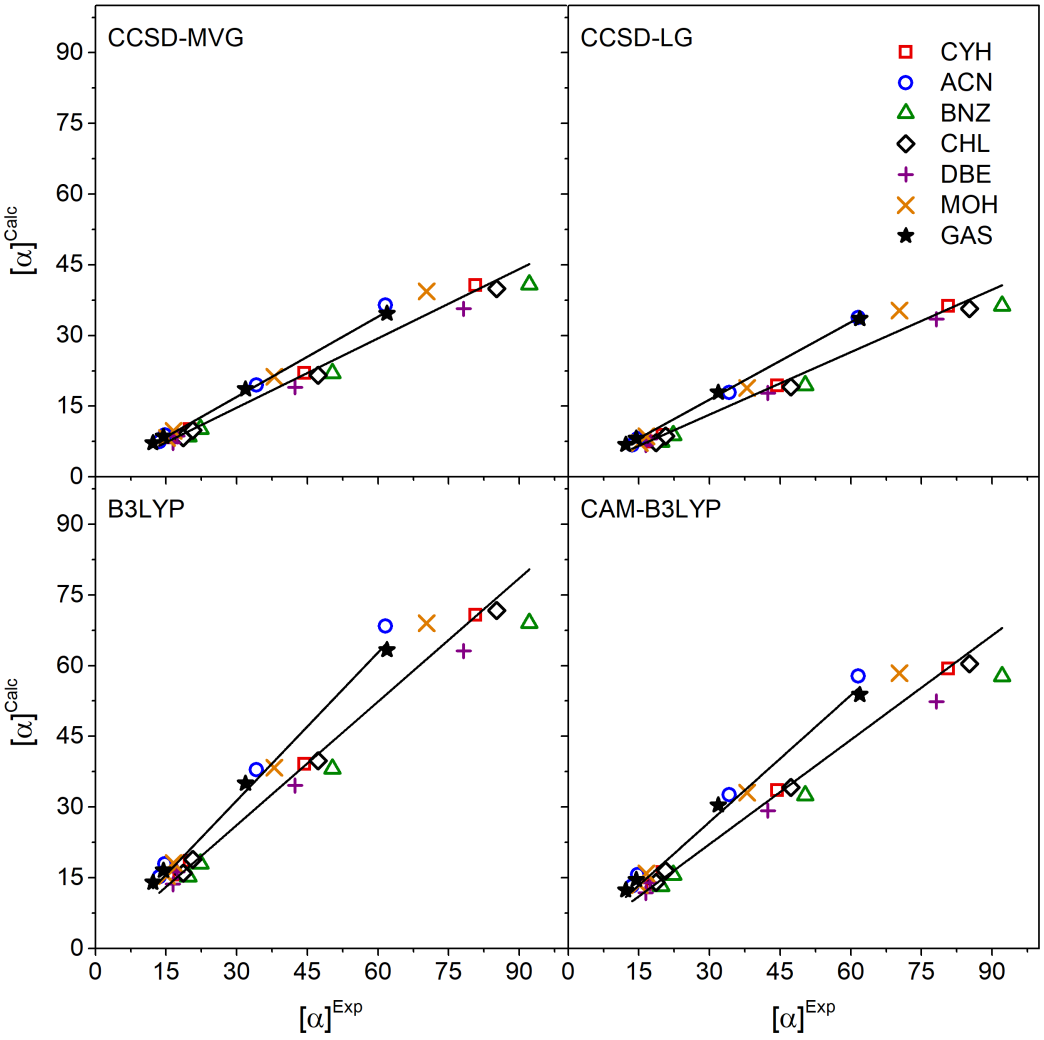


Figure 4.4: Gas and solution phase plots of $[\alpha]^T_\lambda$ (in deg dm^{-1} (g/mL)$^{-1}$) for (*R*)-*cis*-pinane (**3**). The solvent acronyms are listed in Table 1. Experimental results are provided by the Vaccaro Group.

near unity. In solution, experimental values are dispersed in a manner that is not reproduced by the calculations. The solvent-induced perturbation is small in acetonitrile, where the specific rotation is nearly identical to that of the gas phase, and becomes quite large in benzene, where the effective shift represents 50% of the gas-phase results (measured 355 nm values of 61.9, 61.6, and 92.1 deg dm^{-1} $(g/mL)^{-1}$ in gas, ACN, and BNZ). CCSD-MVG predicts that the effect of acetonitrile to be very small and, while the largest change from gas to any solvent still is obtained with benzene, the magnitude of this effect is reduced to only 20% (CCSD-MVG 355 nm values of 34.7, 36.5, and 40.8 deg dm^{-1} $(g/mL)^{-1}$ in gas, ACN, and BNZ). DFT also predicts very small shifts with PCM, amounting to only 10% from gas to benzene in the case of B3LYP (B3LYP 355 nm values of 63.3 and 69.1 deg dm^{-1} $(g/mL)^{-1}$ in gas and BZN). The largest difference in rotatory powers now is between two solvents, chloroform and dibutyl ether, rather than with the gas phase (B3LYP 355nm values of 70.7 and 63.1 deg dm^{-1} $(g/mL)^{-1}$ in CHL and DBE).

Figure 5 illustrates results obtained for compound **4**, which is the only chiral system having a negative optical rotation in the current dataset. Studies performed in benzene are not included, as they seem to exhibit behavior inconsistent with those of other solvents and molecules. Linear correlation plots obtained for the CCSD-MVG gas-phase and solution-phase predictions have nearly identical slopes, which probably is fortuitous given the spread of the experimental results in solution, while other methods yield gas-phase slopes greater than their solution-phase counterparts. The three latter methods also predict acetonitrile rotatory powers to fall essentially along the linear fit of the gas phase. In addition, while the observed range of solvated chiroptical response is appreciable (measured 355 nm values of -45.4 and -64.4 deg dm^{-1} $(g/mL)^{-1}$ in ACN and CYH), there is very little difference between any pair of solvents in the calculations (the largest being found for CCSD-MVG which yields 355nm predictions of -38.7 and -42.7 deg dm^{-1} $(g/mL)^{-1}$ in MOH and CYH). Nevertheless, quantum-chemical analyses reproduce the qualitative trends seen in the experiments, where $[\alpha]_{\lambda}^{T}$ is largest for cyclohexane, followed by dibutyl ether and then methanol, although acetonitrile does not fit this trend. The magnitude of the solution-phase specific rotation increases with increasing ε_{∞} in the calculations, and with decreasing ε_0 in the experiments, but the

two sequences are different, as depicted in Table 1.

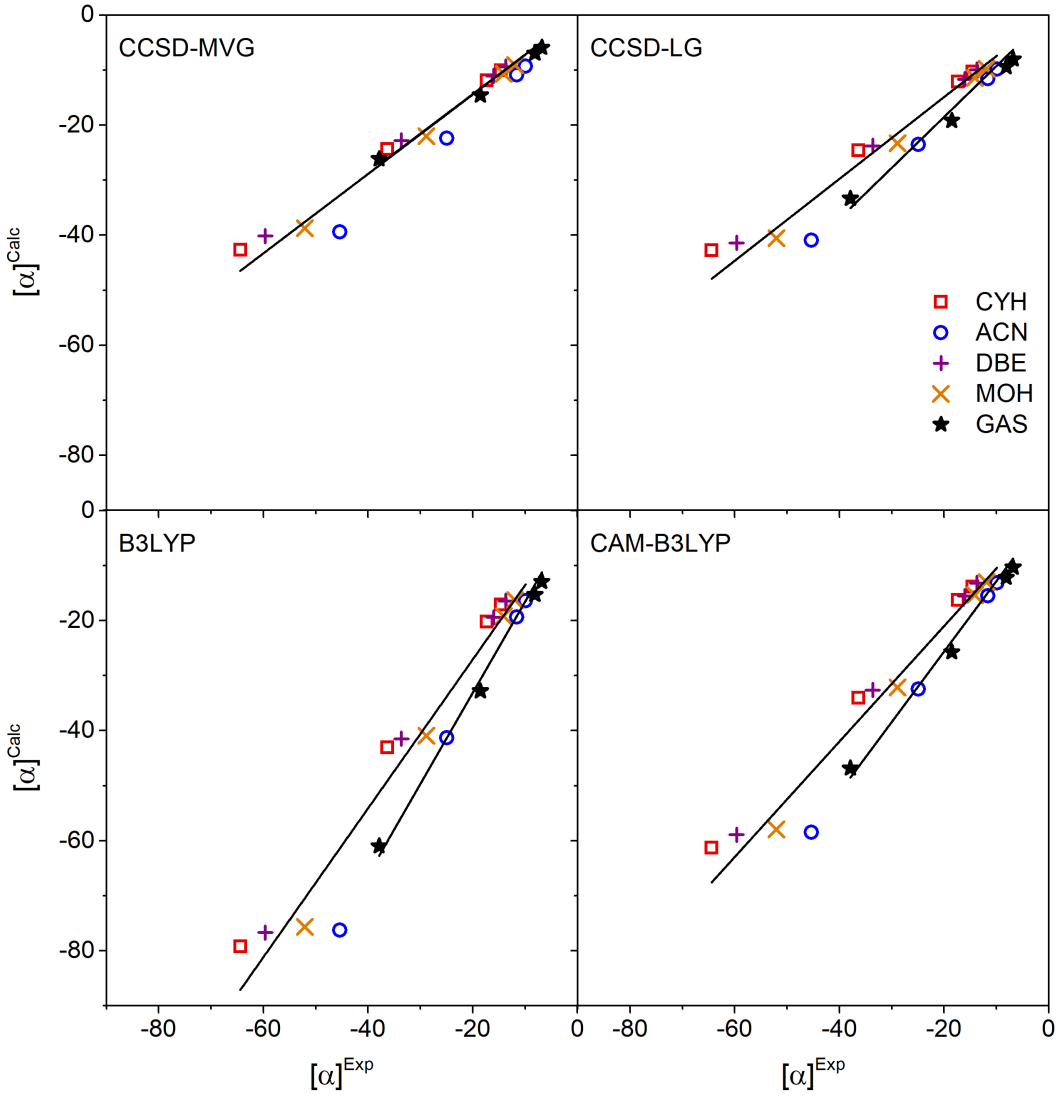


Figure 4.5: Gas and solution phase plots of $[\alpha]^T_\lambda$ (in deg dm^{-1} (g/mL)$^{-1}$) for (*S*)-2-chloropropionitrile (**4**). The solvent acronyms are listed in Table 1. Experimental results are provided by the Vaccaro Group.

We now move to a discussion of solvent shifts, $\Delta[\alpha]^T_\lambda$, the first of which is presented in Figure 6 for molecule **1**. In contrast to other methods, CCSD-MVG successfully reproduces the proper sign

of this quantity for cyclohexane, benzene, and dibutyl ether, but the magnitude is underestimated by a nearly a factor of two. The $\Delta[\alpha]_\lambda^T$ values predicted by CCSD-MVG for methanol also have the same sign as the experimental values, and are reproduced nearly quantitatively as well. The signs of $\Delta[\alpha]_\lambda^T$ estimated by CCSD-MVG for acetonitrile and chloroform are incorrect, although experimental values are difficult to reproduce as their magnitude is smaller than the expected accuracy of the theoretical methods. (experimental {CCSD-MVG} 355 nm values of -1.6 {12.0} and 0.6 {4.9} deg dm^{-1} (g/mL)$^{-1}$ in ACN and CYH). The largest shift both experimentally and theoretically is observed for benzene. All other theoretical approaches predict incorrect signs for the solvent shifts.

The solvent shifts for compound **2** are illustrated in Figure 7. In addition to the over-prediction of solvent effects mentioned above, experiments and calculations also disagree on the direction of the shifts, with the former displaying a negative slope while the latter predict solvation to increase the magnitude of $\Delta[\alpha]_\lambda^T$. Additionally, CCSD-LG and both DFT methods yield very similar $\Delta[\alpha]_\lambda^T$ parameters, while CCSD-MVG suggests much smaller values.

Figure 8 highlights the $\Delta[\alpha]_\lambda^T$ results obtained for compound **3**. The calculations reproduce the sign of the experimental shift (except in the case of acetonitrile), but the magnitude is severely underestimated. Despite an incorrect sign, both choices of gauge for CCSD correctly predict a very small shift in acetonitrile (355 nm shift values of -0.3, 1.8, and 0.2 deg dm^{-1} (g/mL)$^{-1}$ by experiment, CCSD-MVG, and CCSD-LG). The CC methods also duplicate the smaller shift observed in chloroform than in benzene while erroneously giving a greater shift in cyclohexane than in chloroform. The converse is true for DFT, where the acetonitrile results are large (355 nm shift values of 5.1 and 3.9 deg dm^{-1} (g/mL)$^{-1}$ by B3LYP and CAM-B3LYP), the benzene shift is smaller than that of chloroform, and $\Delta[\alpha]_\lambda^T$ in chloroform is larger than in cyclohexane (as found in the experiments).

Figure 9 depicts the $\Delta[\alpha]_\lambda^T$ results for compound **4**. Although the shift magnitudes are underestimated by calculations, particularly by CCSD-LG, the correct sign is obtained. Despite the non-linear trend noted for acetonitrile, which arises from the small experimental change of $\Delta[\alpha]_\lambda^T$

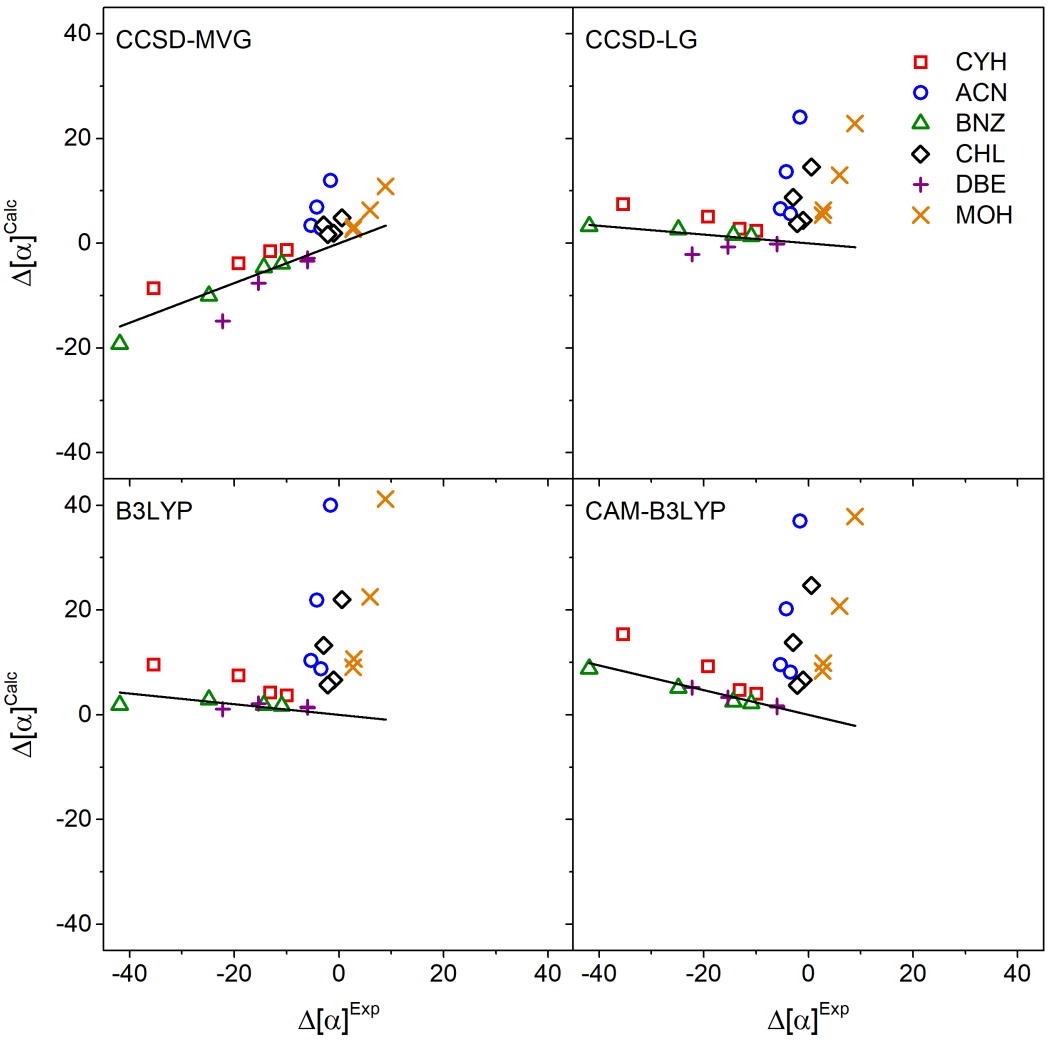


Figure 4.6: The $\Delta[\alpha]^{T}_{\lambda}$ (deg dm^{-1} (g/mL)$^{-1}$) plots for (*R*)-α-pinene (**1**). The solvent acronyms are listed in Table 1. Experimental results are provided by the Vaccaro Group.

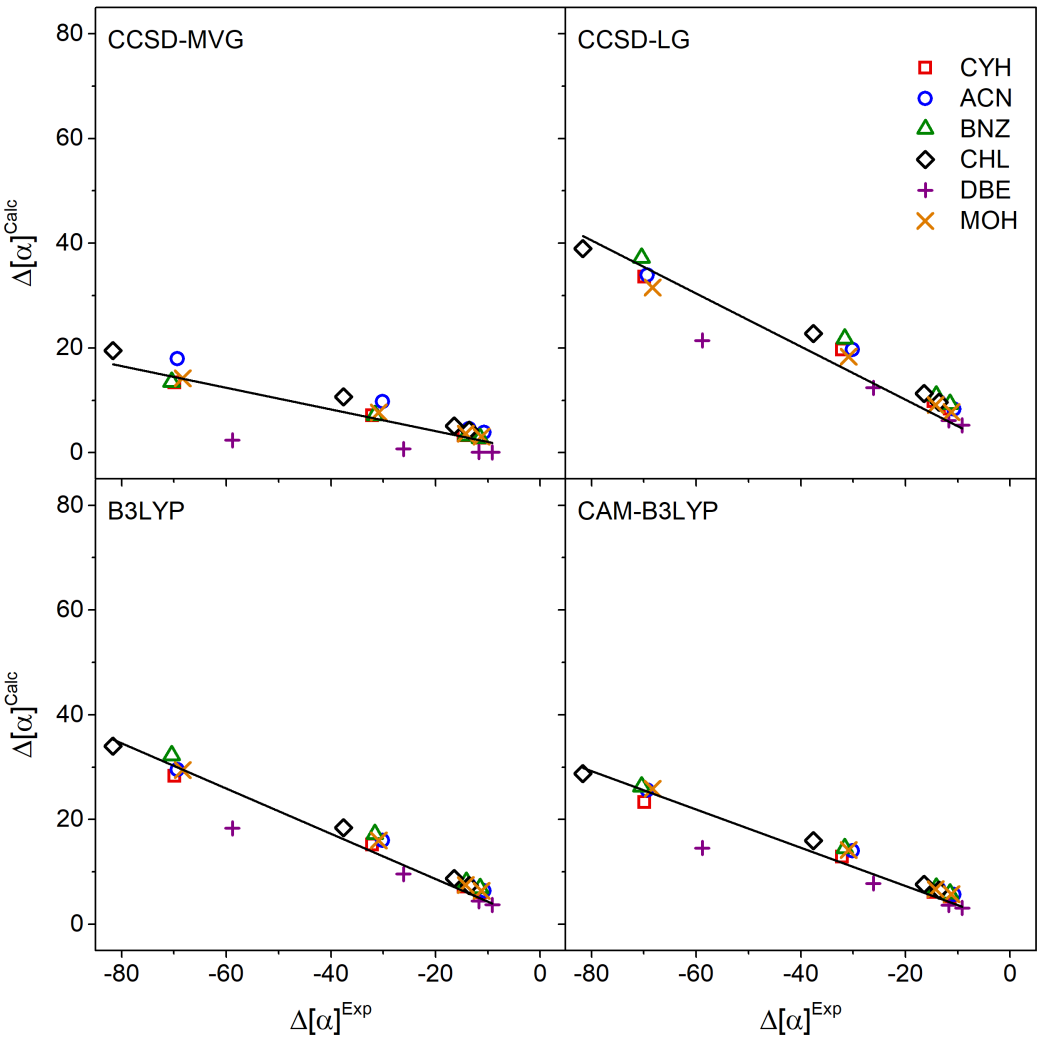


Figure 4.7: The $\Delta[\alpha]^{T}_{\lambda}$ (deg dm^{-1} (g/mL)$^{-1}$) plots for (*S*)-3-carene (**2**). The solvent acronyms are listed in Table 1. Experimental results are provided by the Vaccaro Group.

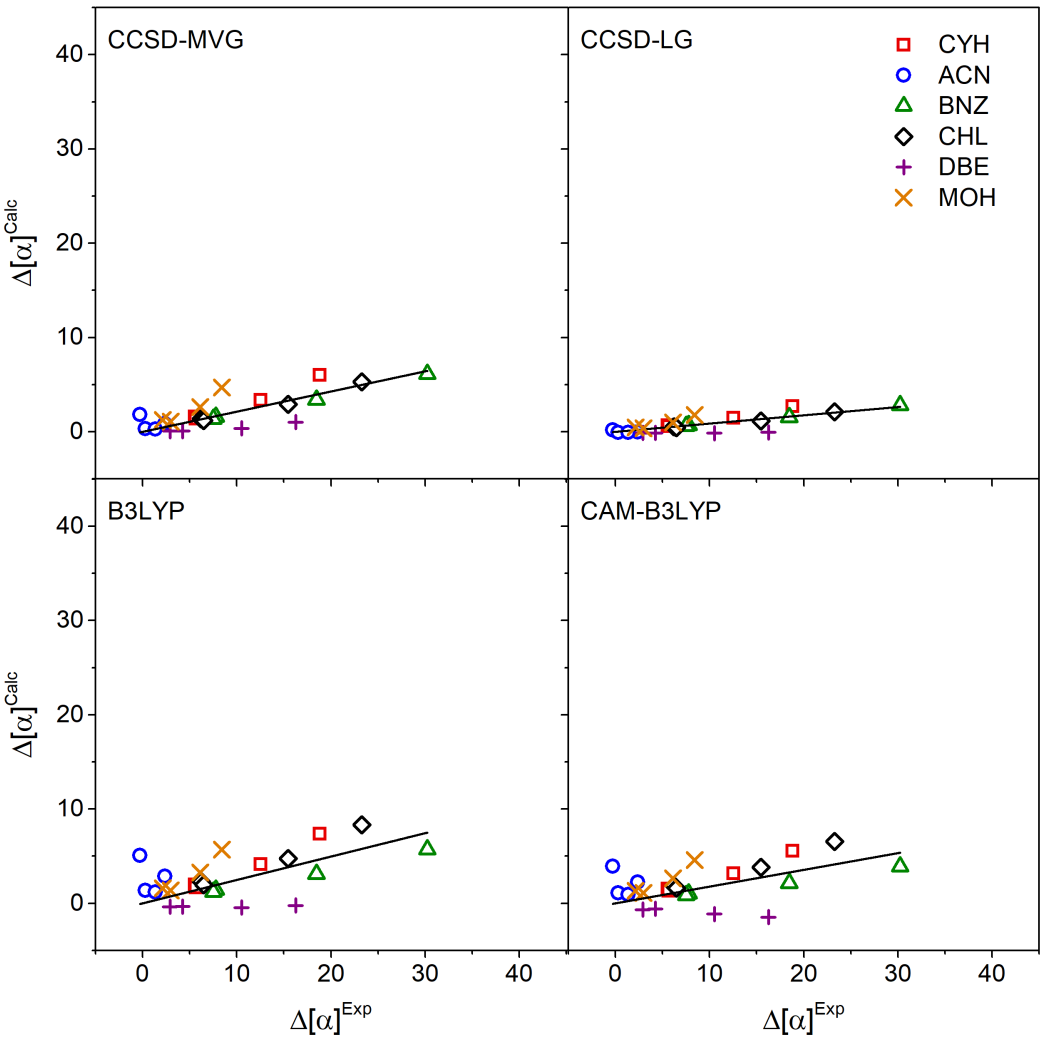


Figure 4.8: The $\Delta[\alpha]^{T}_{\lambda}$ (deg dm^{-1} (g/mL)$^{-1}$) plots for for (*R*)-*cis*-pinane (**3**). The solvent acronyms are listed in Table 1. Experimental results are provided by the Vaccaro Group.

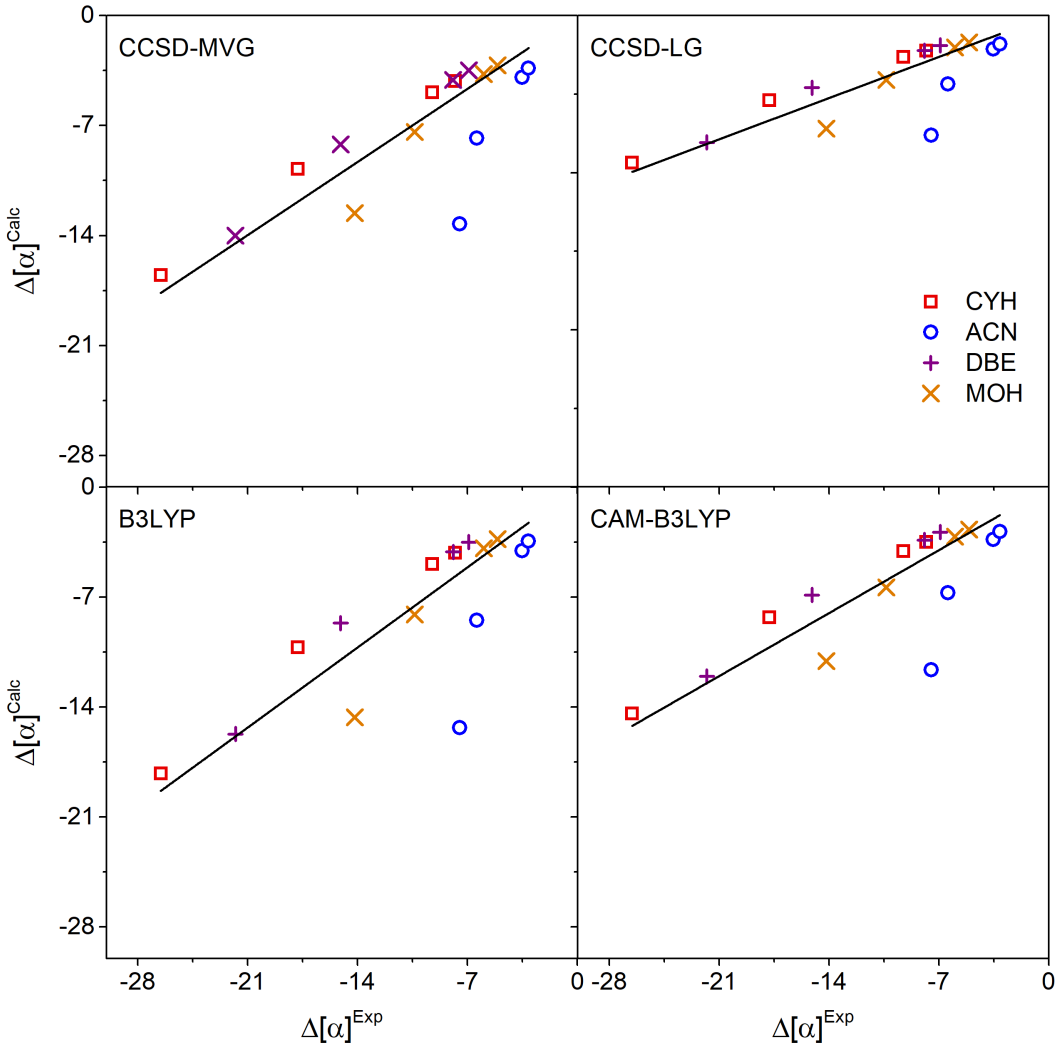


Figure 4.9: The $\Delta[\alpha]^{T}_{\lambda}$ (deg dm^{-1} (g/mL)$^{-1}$) plots for (*S*)-2-chloropropionitrile (**4**). The solvent acronyms are listed in Table 1. Experimental results are provided by the Vaccaro Group.

with wavelength relative to calculated values, calculations for all solvents produce similar shifts. This anomaly may stem from the similar structures of acetonitrile and **4**, which could lead to subtle interactions that differ from those operating in other (bulkier) solvents. Such effects are not reproduced by PCM since the atomistic nature of the solvent is neglected. To a lesser extent, this also is true for methanol, which exhibits comparable behavior. The trend reported for the $\Delta[\alpha]_\lambda^T$ plots of **4** in Figure 5, where the magnitude of the specific rotation in solution increases with increasing ε_∞ in calculations and with decreasing ε_0 in experiments, is present in the $\Delta[\alpha]_\lambda^T$ shifts as well.

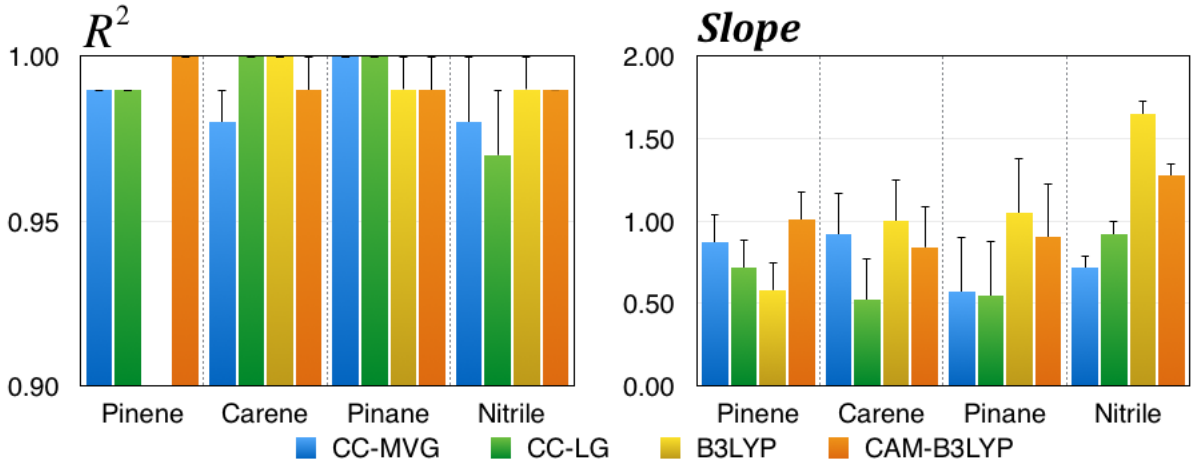


Figure 4.10: The R^2 coefficients and slopes for linear fits of the gas-phase specific rotation. The black bars indicate how the values change when ZPVCs are included.

The trends across molecules can be compared directly by examining the R^2 (correlation coefficient) metrics and attendant slopes extracted from linear regressions in Figures 2-9 (where intercepts were constrained to zero). These results have been compiled in Figures 10 (gas phase) and 11 (solution phase), with the corresponding zero-point vibrational corrections (ZPVCs), as discussed in the Theory section, also being included. Excellent agreement is achieved between experiment and theory in the gas phase. The R^2 values indicate a strong linear correlation to exist between the calculated and experimental rotatory powers, while the accompanying near-unity slopes imply the former to reproduce the latter uniformly. The CCSD-MVG method performs exceptionally well for compounds **1**, **2**, and **4**, but underestimates experiments for **3**. The CAM-B3LYP slopes are close to unity for **1**, **2**, and **3**, suggesting good predictions of polarimetric results, while the

slope for molecule **4** is significantly larger than that for **1**. As shown in Figure 10, vibrational corrections tend to improve calculations by enhancing the quality of correlation attained with measurements. However, because the addition of ZPVCs consistently increases the slope for these systems, analyses that underestimate this quantity are improved while those that do not are made worse. Consequently, the consistently under-predicted CCSD-LG slopes are enhanced towards unity while their over-predicted CAM-B3LYP counterparts always deteriorate upon addition of ZPVCs. The performance of B3LYP and CCSD-MVG with ZPVCs depends on how well experimental values are duplicated initially, although such corrections do tend to improve the overall agreement with experiment.

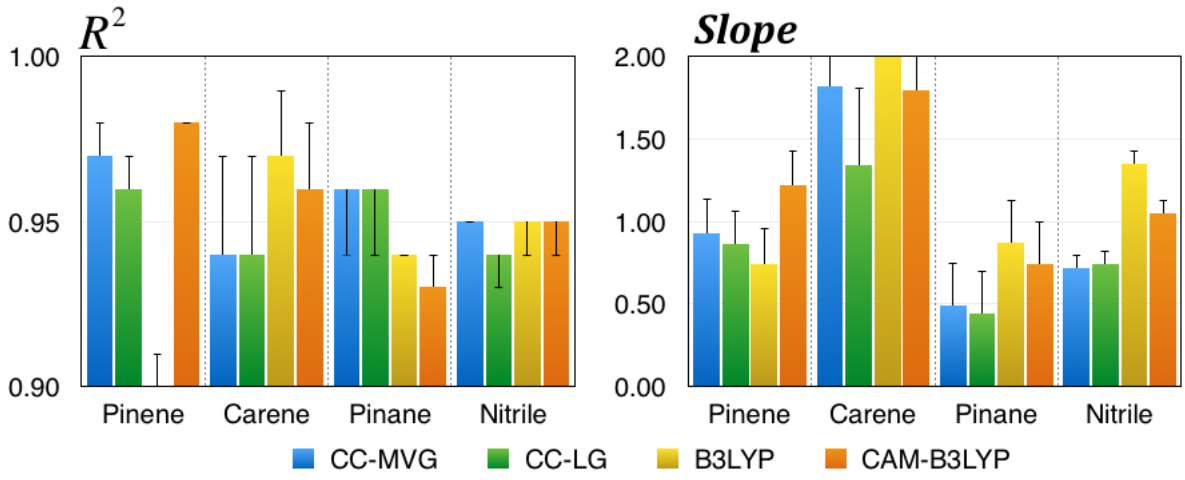


Figure 4.11: The R^2 coefficients and slopes for the linear fits of the solution-phase specific rotation. The black bars indicate how the values change when ZPVCs are included.

The R^2 metrics for solution-phase specific rotations in Figure 11 suggest that the linear correlation between experiment and theory remains reasonable, although not as good as in the gas phase. Indeed, the accompanying slopes imply that solvated results depend strongly on the solute being considered. CCSD methods consistently underestimate experimental findings, except for **2**, which (as mentioned previously) would appear to be an especially difficult case. For compound **1**, slopes from either choice of gauge for CCSD are just under unity, while those for **3** and **4** are much smaller than unity. DFT methods have similar difficulties predicting $[\alpha]^T_\lambda$ for **2**, but also tend to overestimate or underestimate rotatory powers for the remaining three molecules without

a clear pattern. ZPVCs improve the quality of correlation attained for **1** and **2**, while worsening it for the other two species of interest. Such corrections tend to increase the slope of the linear fit obtained for all methods and molecules, thus (as found in the gas-phase) enhancing the performance of methods that under-predict experiments while diminishing that of their over-predicting counterparts.

Figure 12 contains the R^2 (correlation-coefficient) metrics and slopes extracted from consideration of solvent shifts, both of which suggest the correlation between experiment and theory to be exceptionally weak. The R^2 parameters for plots of **2** are much better than those of the other solutes, yet calculations predict corresponding shifts in the wrong direction. The addition of ZPVCs strengthens the correlations for **2** and **3**, but weakens them for **1** and **4**. Slopes are underestimated for every molecule, although only those for **3** and **4** have the correct sign, as does that for **1** evaluated by CCSD-MVG. All slopes for the shifts of **2** have the wrong sign, as do those of **1** for almost all methods. The addition of ZPVCs increases the slope for **3** and **4**, but the solvent shift remains under-predicted. For **1** and **2**, where calculations predict the incorrect direction of the shift, ZPVCs further increase the slope magnitude towards more negative values. This indicates that such vibrational corrections do not produce a general improvement of calculated results for the present set of target molecules.

4.4 Conclusion

A detailed study has been performed to elucidate the influence of solvation effects on the dispersive optical activity of four small organic molecules: (*R*)-α-pinene (**1**), (*S*)-3-carene (**2**), (*R*)-*cis*-pinane (**3**), and (*S*)-2-chloropropionitrile (**4**). In particular, this test set was exploited to systematically examine the quality of agreement attained between experimental measurements of specific rotation, $[\alpha]^T_\lambda$, and calculations performed with the polarizable continuum model (PCM) and three levels of theory: B3LYP, CAM-B3LYP, and CCSD, where the latter employed two choices of gauge for the magnetic perturbation. In the gas phase, theory and experiment display excellent correlations, in keeping with previous reports.[140] Indeed, despite specific cases of overestimated

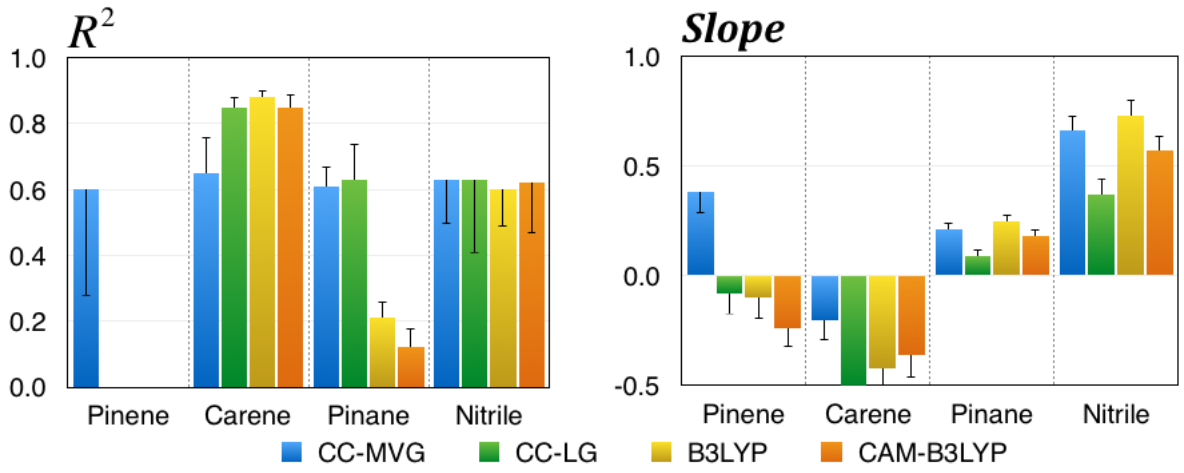


Figure 4.12: The R^2 coefficients and slopes for the linear fits of the solvent shifts. The black bars indicate how the values change when ZPVCs are included. The R^2 values for (*R*)-α-pinene are negative (not reported), which is indicative of very weak correlation between data points.

or underestimated predictions, the relationship between theoretical and experimental rotatory powers (*viz.*, the experiment-theory correlation) in the gas phase is strongly linear for all methods (the only exception being B3LYP results obtained for molecule **1**). Most importantly, all quantum-chemical analyses correctly reproduce the sign of the observed chiroptical response. The addition of zero-point vibrational corrections tends to increase the slope of the linear fit of $[\alpha]_\omega$, favoring CC results, which underestimate experimental findings, while the agreement using DFT become worse, as DFT tends to overestimate experimental results.

Experiments clearly show solvation can have a large effect on dispersive optical activity, with each solute-solvent pair behaving differently. The PCM does predict an overall effect on $[\alpha]^T_\lambda$, but its magnitude is significantly underestimated. This trend is apparent in Figures 4 and 5, where there is a large disparity among rotatory powers measured in solution, while the corresponding spread for calculations is minimal. CCSD with both choices of gauge underestimates the specific rotation in solution for three of the four target molecules, while overestimating the effect for **2**. On the other hand, the various functionals considered by the present work provide a somewhat mixed behavior (*cf.* Figures 10 and 11), as the results of DFT calculations are system and property dependent. For all methods, the linear-fit correlation found in the gas phase seems to deteriorate in the presence of

a solvent medium. The addition of ZPVCs has a similar effect to that in the gas phase, causing the underestimated predictions for **3** and **4** to be improved while the kindred overestimated results for **1** and **2** are made worse.

Solvation shifts (*cf.* Eq. 7) are found to be very difficult to reproduce, particularly for those cases where the magnitude of $[\alpha]_\lambda^T$ decreases in solution. Such behavior is observed for compounds **1** and **2**, and PCM is unable to reproduce these effects except in the case of **1** with CCSD-MVG. Other solution-phase calculations always give larger rotatory powers than those predicted for the gas phase, independent of the initial sign or direction of the experimental solvent shift. For molecules **3** and **4**, ZPVCs produce a minute increase towards experimental values of the solvent shift, whereas for **1** and **2** the solvent shifts tend to worsen upon incorporation of such vibrational corrections.

The emerging results suggest that it may be better to compute the value of the specific rotation in solution directly rather than attempt to obtain solvation shifts with PCM, as one may apply empirical corrections to account for any systematic offsets. However, the attendant error seems more molecule dependent than in gas phase, as demonstrated succinctly in the case of (*S*)-3-carene (**2**). The shortcomings in PCM may be due to the fact that the current model is purely electrostatic in nature, and the missing effects of dispersion and repulsion may play an important role for chiroptical properties.[146] Additionally, the process of "chiral imprinting", as embodied in the chiral ordering induced in the achiral solvent molecules that constitute the first solvation shell, has been shown to afford significant contributions to $[\alpha]_\lambda^T$ in certain cases, and such phenomena cannot be reproduced by PCM.[57] Thus, dispersive optical-activity calculations in solution based upon the standard PCM methodology do not seem to offer the same predictive power as kindred calculations performed in the isolated-molecule limit of the vapor phase, and more sophisticated treatments of solvation appear to be necessary.

4.5 Acknowledgements

T.A. and M.C. are grateful for support from the National Science Foundation under CHE-1650942. P.M.L. and P.H.V. gratefully acknowledge the support of the U.S. National Science Foundation under the auspices of Grant No. CHE-1464957.

4.6 Supporting Information

The supporting information contains the optimized geometry of every molecule in every solvent; the calculated, measured, and extrapolated values of the specific rotation at each wavelength and in each phase, the corresponding shifts, and the extrapolation parameters used in Eq. 6 are also reported.

Chapter 5

A Molecular Orbital Selection Approach for Fast Calculations of Specific Rotation with Density Functional Theory

5.1 Introduction

The study of optically active molecules is of great interest due to the fundamental role they play as the building blocks of life (L-amino acids, D-sugars), and their importance in the pharmaceutical industry.[2] Due to their pharmaceutical relevance, facile assignment of the absolute configuration (AC) of chiral molecules is important because it is directly tied to their bio-active properties. The ability to directly link a structure to the chiroptical property of interest, such as optical rotation (OR), led electronic structure calculations to become a standard for identifying the AC of chiral molecules. Accurate calculations require the proper account of electron correlation, and large basis sets.[2] To this end, great progress has been made in recent years in applying both density functional theory (DFT) and coupled cluster (CC) methods to the calculation of $[\alpha]_\omega$.[5–8,10–20,147] However, as systems of interest become larger, the cost of electronic structure calculations becomes prohibitive, and compromises between cost and accuracy are needed.[140]

Despite the increasing complexity of systems of interest, few attempts have been made to reduce the cost of $[\alpha]_\omega$ calculations as a whole. Wiberg *et al.* calculated $[\alpha]_\omega$ with small basis sets (STO-3G and 3-21G) augmented with diffuse functions from large correlation consistent basis sets. These mixed basis sets reproduce the values obtained with the full basis sets at a fraction of the computational cost.[61] The optical rotation prediction (ORP) basis set, developed specifically to calculate $[\alpha]_\omega$, is able to produce aug-cc-pVTZ level quality but is closer in size to aug-cc-pVDZ.[62,63]

Crawford *et al.* extended the local correlation idea of Pulay and Saebø[148,149] to response properties,[150] and then later implemented it for $[\alpha]_\omega$ calculations.[64] This method uses localized orbitals, and neglects interactions between distant orbitals, thus reducing the cost of the correlation part of the calculation.[65,66] While this method is successful for smaller systems, the authors show that for larger systems the need for tighter thresholds outweighs the benefits of neglecting parts of the wave function.[64]

In this work, we present a different approach to reduce the cost of $[\alpha]_\omega$ calculations with Kohn-Sham density functional theory (KS-DFT), based on the selection of the molecular orbitals (MOs) that are likely to contribute the most to this property, and discarding the rest. We have recently shown that a significant portion of the MOs do not contribute significantly to $[\alpha]_\omega$ through a post-calculation analysis of the optical rotation tensor.[23,26,98] Here, we try to determine the meaningful MOs beforehand, and to solve the linear response equations only with a subset of relevant MOs while preserving accuracy.

The paper is organized as follows: Section 5.2 outlines the procedure for the orbital selection, section 5.3 contains the results of the calculations, and section 5.4 presents a discussion and concluding remarks.

5.2 Theory and Computational Details

In order to explain how the orbital selection process works, it is useful to briefly review how the specific rotation is typically computed. For molecules in isotropic media, the specific rotation is related to trace of the electric dipole-magnetic dipole polarizability tensor $G'_{\alpha\beta}$:[91]

$$\beta(\omega) = -\frac{1}{3}\sum_{\alpha} G'_{\alpha\alpha} \tag{5.1}$$

where ω is the frequency of incident light, and the tensor $G'_{\alpha\beta}$ is defined as:

$$G'_{\alpha\beta} = -2\omega \sum_{n\neq 0} \mathrm{Im}\frac{\langle\psi_0|\mu_\alpha|\psi_n\rangle\langle\psi_n|m_\beta|\psi_0\rangle}{\omega_{n0}^2 - \omega^2} \tag{5.2}$$

where μ is the electric dipole moment operator, m is the magnetic dipole moment operator, the Greek letter indexes represent Cartesian coordinates, and ψ_0 and ψ_n are the ground and excited electronic state wave functions, respectively. However, this series is slowly converging, and evaluating electronic excited states is computationally expensive. A significantly more efficient approach is based on time-averaged variational linear response theory,[19,93] where the evaluation of the transition moments in Eq. 5.2 is replaced by the evaluation of a perturbed electron density, P^x_α, where x is the external field perturbation. Using Hartree-Fock (HF) or KS-DFT as approximations of the unperturbed density, and using the variational nature of these methods, the perturbed density can be obtained by solving the coupled perturbed HF or KS equations (CPHF or CPKS):[19,151]

$$\left[\begin{pmatrix} \mathbf{M} & \mathbf{Q} \\ \mathbf{Q}^* & \mathbf{M}^* \end{pmatrix} - \omega \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix}\right] \begin{pmatrix} \mathbf{P}^S_\beta \\ (\mathbf{P}^S_\beta)^* \end{pmatrix} = \begin{pmatrix} \mathbf{S}_\beta \\ \mathbf{S}^*_\beta \end{pmatrix} \tag{5.3}$$

where $M_{ia,jb} = (\varepsilon_a - \varepsilon_i)\delta_{ij}\delta_{ab} + K_{aj,ib}$, $Q_{ia,jb} = K_{ab,ij}$, $K_{aj,ib} = \langle aj|ib\rangle + \langle aj|f^{\mathrm{XC}}(r,r')|ib\rangle$, and $\mathbf{1}$ is the unit matrix. Extra terms should appear in Eq. 5.3 when perturbation-dependent atomic orbitals (such as gauge including atomic orbitals, or GIAOs[19,94,95]) are used, but they are neglected here for simplicity. The indexes i, j denote occupied MOs, a, b virtual MOs, ε are the MO energies, $\langle aj|ib\rangle$ are two electron repulsion integrals (2ERIs), and $f^{\mathrm{XC}}(r,r')$ is the exchange correlation kernel of a density functional. In Eq. 5.3, the matrices $S_{\beta,ai} = \langle a|S_\beta|i\rangle$ and their complex conjugate represent the perturbation integrals in MO basis: $\mathbf{S}_\beta = \mu_\beta$ or $\mathbf{m}_\beta$. The G' tensor is then obtained by tracing the perturbed density with the conjugate dipole integrals:

$$G'_{\alpha\beta} = -\omega \mathrm{Tr}(\mathrm{Im}\{\mathbf{R}_\alpha(\mathbf{P}^S_\beta)^* + c.c.\}) \tag{5.4}$$

where $\mathbf{R}_\alpha = \mathbf{m}_\alpha$ or μ_α.

Eq. 5.3 is generally not solved by direct inversion of the matrix on the left-hand side because the transformation of the 2ERIs to MO basis is computationally expensive (it scales as $O(N^5)$, where N is the basis set size), and the matrix itself is rather large (i.e., the dimension is $N_{occ} * N_{vir}$, where

N_{occ} is the number of occupied MOs and N_{vir} is the number of virtual MOs). Instead, efficient iterative algorithms based on partial AO $\leftrightarrow$ MO basis transformations are employed.[152,153] These algorithms avoid the full 2ERIs transformation and the storage of the **M** and **Q** matrices. The scaling is thus reduced formally to $O(N^4)$, but practically to $O(N^3)$ by using efficient integral evaluation and contraction algorithms.

In the context of this work, the key point is that any iterative algorithm generates a guess for the solution vectors, see Eq. (64) in Ref. 152, to start the solution of Eq. 5.3. Thus, we can use the guess $\mathbf{P}^S$ to define two possible criteria for the MO selection, one based on the square elements of the guess matrix:

$$G^1_{ia} \equiv \sum_{\beta} \left| P^S_{\beta,ia} \right|^2 \tag{5.5}$$

and one based on the dot product with the conjugate dipole integrals:

$$G^2_{ia} \equiv \tilde{S}_{ia} = \mathrm{Im} \left\{ \sum_{\beta} R_{\beta,ai} (P^S_{\beta,ia})^* + c.c. \right\} \tag{5.6}$$

where we have used the $\tilde{S}$ symbol to connect this quantity with the rotatory strength in configuration space that we defined for the qualitative analysis of the MO transition contributions to the OR tensor.[23,26,98] The difference here is that the $\tilde{S}$ values are computed with the guess density rather than the converged density. In particular, we solve the CPKS equations for the magnetic dipole perturbation, so that $S = m$ and $R = \mu$ in Eqs. 5.5 and 5.6. Therefore, an occupied MO is selected if:

$$|G^1_{ia}| \;\text{ or }\; |G^2_{ia}| > \varepsilon \quad \forall a \tag{5.7}$$

and a virtual MO is selected if:

$$|G^1_{ia}| \;\text{ or }\; |G^2_{ia}| > \varepsilon \quad \forall i \tag{5.8}$$

depending on the choice of selection criterion. The threshold ε is defined either as an absolute value, e.g., $\varepsilon = 10^{-n}$ where n is a positive integer, or relative to the largest value of $|G_{ia}|$, e.g., $\varepsilon = 10^{-n}\mathrm{Max}\{|G_{ia}|\}$. In the following section, we test these four possibilities for the MO selection

criterion with three values of n: $n = 3,\ 4,\ 5$.

Since the scope of this work is only to test the validity of this idea, we remove the discarded MOs from the calculation simply by setting to zero the corresponding unperturbed MO coefficients in the CPKS calculation. This ensures that these MOs do not contribute to the evaluation of the OR tensor, but it does not effectively change the cost of the calculation. Instead, we estimate the potential cost savings by assuming a cubic scaling with the size of the basis set, as mentioned above, and reporting a speed-up defined as:

$$\text{Speed-Up} = (N_{tot}/N_{sel})^3 \tag{5.9}$$

where N_{tot} is the total number of MOs, and N_{sel} is the number of selected MOs.

All calculations were performed with a development version of the GAUSSIAN suite of programs.[114] Geometries were optimized using the CAM-B3LYP/aug-cc-pVDZ model chemistry.[115,116] Calculations of specific rotation were performed at the sodium-D line ($\omega = 589.3$ nm) with the B3LYP[117–119] and CAM-B3LYP functionals, and the aug-cc-pVDZ and aug-cc-pVTZ basis sets in the length gauge formalism using GIAOs[19,94,95].

5.3 Results

To test the success of the selection procedure outlined in section 5.2, we performed calculations using a total of 51 molecules. This set is constructed from the 42 organic molecules provided in the OR45 test set of Srebro *et al.*,[140] plus 7 more organic molecules shown in Figure 5.1: two conformers of 2-carene (**43** and **44**), two conformers of 3-methylcyclopentanone (**45** and **46**), cycloserine (**47**), fucose (**48**), limonene (**49**), nicotine (**50**), and oxaceprol (**51**). The absolute values of the specific rotation range across 5 orders of magnitude ($10^0 - 10^4$ deg dm^{-1} (g/mL)$^{-1}$), reported in Table 1 of the Appendix, providing a stringent test for the procedure.

We define a short-hand notation to distinguish the choice of selection criterion: A/R represents absolute or relative thresholds, M/S represents G^1 or G^2 (in Eqs. 5.5-5.6), respectively, followed

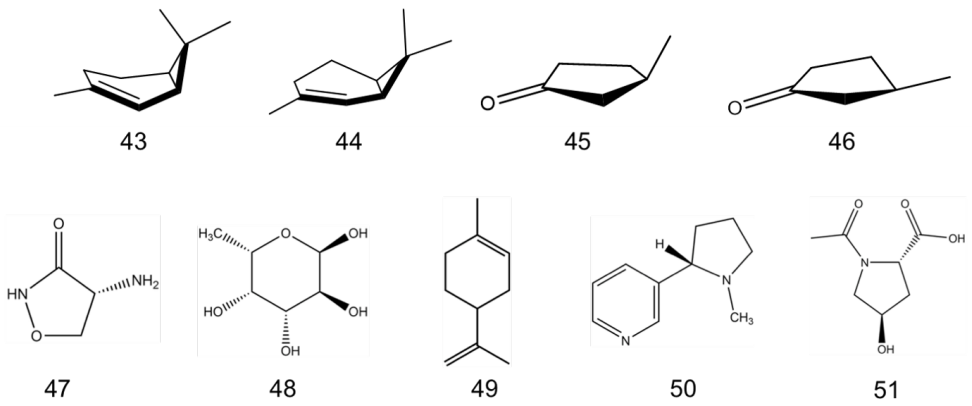


Figure 5.1: Structures of molecules **43–51**.

by the threshold order of magnitude (n). For instance, AM5 indicates the G^1 criterion with the absolute threshold $\varepsilon = 10^{-5}$. The results are presented in terms of relative errors, with histograms that collect the error distribution for the test set, and tables that include the mean signed error (MSE), mean unsigned error (MUE), maximum error (Max), the MSE standard deviation (σ_{MSE}), the average relative number of selected MOs ($\bar{N}_{sel}$), its standard deviation ($\sigma_{\bar{N}_{sel}}$), and the estimated speed-up based on Eq. 5.9. All values of $[\alpha]_D$ with all selection criteria are reported in Tables 1-8 of the Appendix.

The relative error distribution for CAM-B3LYP/aug-cc-pVDZ and the AGn criterion is reported in Figure 5.2. The plots for the other three selection criteria are similar, and are reported in Figures 1-3 of the Appendix. The collective statistical analysis for this model chemistry and all selection criteria is reported in Table 5.1. It is immediately apparent from the figure that the error distribution for AG5 is sharply centered around 0, and it spreads out as n decreases. Similar trends are obtained with the other criteria. All calculations with thresholds $n > 3$ reproduce $[\alpha]_D$ very well, as indicated by both the small MUE and σ_{MSE} in Table 5.1. The AG5 results are very accurate, with a MUE of only 0.8%, a σ_{MSE} of 1.4%, and a Max of only 5.5%. This selection criterion discards about 40% of the MOs, with an estimated speed-up of 4x, which is already significant. Using $n = 4$ leads to a slight drop in accuracy, but MSE and MUE are still small, -1.1% and 3.9%, respectively. The values of σ_{MSE} and Max are larger, 8.5% and 49%, respectively. However, this is due to molecules with small absolute values of $[\alpha]_D$, where small absolute errors may result in seemingly

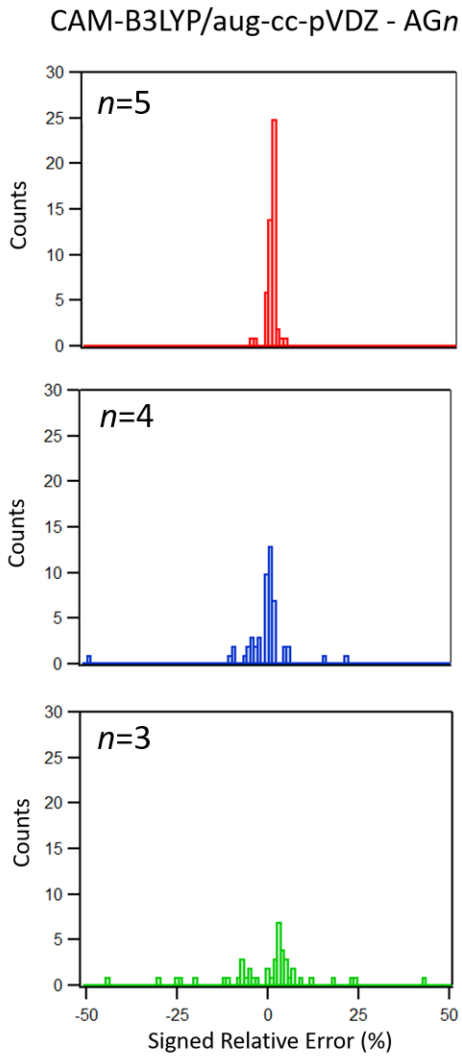


Figure 5.2: Histogram of the signed relative errors for each of the 51 molecules using the AG*n* selection criterion and the CAM-B3LYP/aug-cc-pVDZ method.

Table 5.1: Relative (%) error: MSE, MUE, Max, and σ_{MSE}, average fraction (%) of selected MOs ($\bar{N}_{sel}$) and the corresponding standard deviation ($\sigma_{\bar{N}_{sel}}$), and estimate of the average speed-up (S.U., see Eq. 5.9) for calculations with the aug-cc-pVDZ basis set.

	CAM-B3LYP			B3LYP		
n	5	4	3	5	4	3
			AGn			
MSE	-0.1	-1.1	-1.7	-0.2	-0.3	-7.8
MUE	0.8	3.9	23	0.7	3.5	21
Max	5.5	49	211	7.3	33	228
σ_{MSE}	1.4	8.5	51	1.5	7.7	47
$\bar{N}_{sel}$	62	45	24	63	46	25
$\sigma_{\bar{N}_{sel}}$	5	5	4	5	5	3
S.U.	4.1	11	70	4.0	11	60
			RGn			
MSE	0.0	0.6	-0.7	0.1	0.1	-5.3
MUE	0.2	1.4	6.1	0.7	2.4	17
Max	1.7	13	28	9.4	21	161
σ_{MSE}	0.4	2.6	10	1.9	4.7	37
$\bar{N}_{sel}$	72	57	37	63	47	27
$\sigma_{\bar{N}_{sel}}$	10	12	11	15	14	12
S.U.	2.6	5.5	20	4.0	10	50
			ASn			
MSE	0.0	-0.6	-4.4	0.0	-0.1	-10.0
MUE	0.2	1.9	16.3	0.2	2.0	18
Max	1.2	9.5	170	1.5	18	200
σ_{MSE}	0.3	3.2	36	0.4	4.4	43
$\bar{N}_{sel}$	73	54	28	72	56	29
$\sigma_{\bar{N}_{sel}}$	10	5	5	9	5	5
S.U.	2.6	6	43	2.6	6	42
			RSn			
MSE	0.1	0.1	0.2	0.0	0.1	0.3
MUE	0.2	0.2	1.9	0.1	0.2	2.0
Max	1.7	1.5	9.3	1.1	0.9	11
σ_{MSE}	0.4	0.4	2.9	0.3	0.3	3.3
$\bar{N}_{sel}$	79	71	55	75	67	49
$\sigma_{\bar{N}_{sel}}$	11	12	12	11	13	14
S.U.	2.0	2.8	6.2	2.4	3.3	8.3

large relative errors. For instance, the Max = 49% comes from molecule **1**, which has $[\alpha]_D$ = -19 deg dm^{-1} (g/mL)$^{-1}$, and an absolute error of -9 deg dm^{-1} (g/mL)$^{-1}$; note that both the $[\alpha]_D$ value and the error are considerably below the average expected error for this choice of model chemistry,

i.e., 25-30 deg dm^{-1} (g/mL)$^{-1}$.[140] The other two molecules with seemingly large relative error in Figure 5.2 for AG4 are **19**, which has $[\alpha]_D$ = 10 deg dm^{-1} (g/mL)$^{-1}$ and an absolute error of 2.1 deg dm^{-1} (g/mL)$^{-1}$, and **6**, with $[\alpha]_D$ = 15 deg dm^{-1} (g/mL)$^{-1}$ and an absolute error of 2.4 deg dm^{-1} (g/mL)$^{-1}$. Therefore, the AG4 criterion does in fact provide even better results than what appears in Table 5.1 if the statistical analysis had been performed subdividing the test set in compounds with large and small $[\alpha]_D$, and the absolute rather than the relative error had been used for the latter set. On average, this selection criterion eliminates more than half of the MOs (55%), with a considerable estimated speed-up of 11x. On the other hand, the AG3 criterion is too loose, as shown both in Figure 5.2 and in Table 5.1, and the large speed-up comes at the price of poor accuracy.

The error distributions for the RG*n* criterion at CAM-B3LYP/aug-cc-pVDZ level are shown in Figure 1, and the statistical data are reported in Table 5.1. This criterion provides essentially the same trends as the corresponding AG*n* choice. However, the error averages and spread tend to be smaller than those for AG*n* for the same level of n . This is because a smaller number of MOs (about 10% less) are discarded with RG*n* than with AG*n* for every n . The better accuracy comes with smaller speed-up by a factor of ~ 2. The results for the AS*n* criterion are reported in Figure 2 and Table 5.1. Calculations performed with this criterion provide results that are very close to those with RG*n* in terms of accuracy, number of discarded MOs, and corresponding speed-up. Finally, the RS*n* data, reported in Figure 3 and Table 5.1, are virtually identical to those obtained with the RG($n+1$) criterion. Also for these criteria, the largest relative errors are due to molecules with small $[\alpha]_D$, for instance, the Max error for RG4 (13%) is due to molecule **18**, with $[\alpha]_D$ = 2.3 deg dm^{-1} (g/mL)$^{-1}$ and an error of 0.29 deg dm^{-1} (g/mL)$^{-1}$.

The error distributions with the B3LYP/aug-cc-pVDZ model chemistry and the AG*n* criterion are shown in Figure 5.3, while similar plots for the other criteria are shown in Figures 4-6 of the Appendix. The statistical analysis for all criteria is also reported in Table 5.1. The qualitative trends in the error distribution are very similar to those for the CAM-B3LYP functional. The statistical data also mirror that of CAM-B3LYP, both in terms of accuracy and computational cost. The only

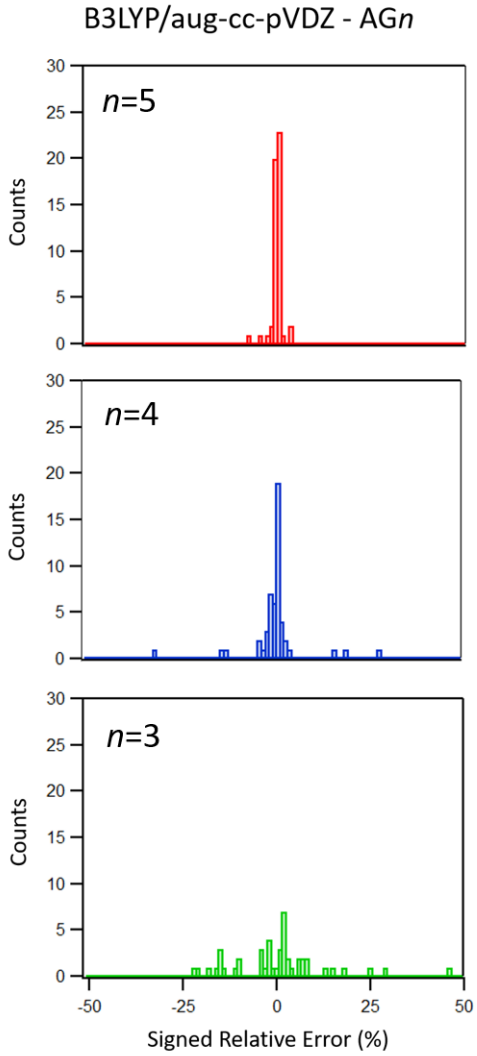


Figure 5.3: Histogram of the signed relative errors for each of the 51 molecules using the AG*n* selection criterion and the B3LYP/aug-cc-pVDZ method.

notable difference is that the errors with the RG*n* criterion are larger than those for CAM-B3LYP because a larger number of MOs are discarded. A similar behavior, albeit to a smaller extent, is obtained with the RS*n* criterion. This difference is due to the fact that the magnitude of the larger G^1_{ia} (and G^2_{ia}) values is larger with CAM-B3LYP than with B3LYP, which results in a looser selection threshold for the latter method. Six molecules in the AG4 set have errors greater than 10%: **1**, **6**, **9**, **18**, **19**, and **34**, with errors of 33%, 28%, 15%, 19%, 16%, and 14%, respectively. The $[\alpha]_D$ for these molecules are: -17, 12, 10, 1.9, 9, and -9 deg dm^{-1} (g/mL)$^{-1}$, respectively, while the errors are 5.6, 3.24, 1.46, 0.35, 1.34, and 1.25 deg dm^{-1} (g/mL)$^{-1}$, respectively. With RG4, molecules **9**, **18**, and **34** have a relative error greater than 10%: -15%, -15%, and 21%, corresponding to absolute values of 1.46, 0.27, and 1.86 deg dm^{-1} (g/mL)$^{-1}$, respectively, for $[\alpha]_D$ values of -10, 1.9, and -9 deg dm^{-1} (g/mL)$^{-1}$, respectively. The B3LYP AS4 set contains only three molecules with errors greater than 10% (**1**, **9**, and **18**), where the errors are 17%, 18%, and 11%, respectively, and the corresponding $[\alpha]_D$ values are also small: -17, -10, and 1.90 deg dm^{-1} (g/mL)$^{-1}$, respectively. No errors greater than 10% are found for the RS4 criterion. These values of $[\alpha]_D$ and corresponding errors are again below the expected accuracy of this level of theory, estimated to be in the 20-25 deg dm^{-1} (g/mL)$^{-1}$ range.[140,154]

The error distribution for CAM-B3LYP and B3LYP with the aug-cc-pVTZ basis set and the AG*n* criterion are shown in Figures 5.4 and 5.5, while the same plots with the other selection criteria are reported in Figures 7-12 in the Appendix. The statistical data of the error and computational cost are shown in Table 5.2. The qualitative trends of error distribution with this basis set are similar to those with aug-cc-pVDZ, except that the distribution is now slightly more spread out. The statistical data in the table for CAM-B3LYP show that, for each selection criterion, there seems to be a correspondence with the $n-1$ choice of the aug-cc-pVDZ basis set in terms of accuracy, amount of selected MOs, and speed-up. This is related to the fact that with a larger basis set there are more virtual MOs available, corresponding to a larger number of $P^S_{\beta,ai}$ elements. More of these elements are small compared to the aug-cc-pVDZ case, but they still contribute significantly to the final value of the specific rotation. Nonetheless, a tight choice of n ($n = 5$) for aug-cc-pVTZ

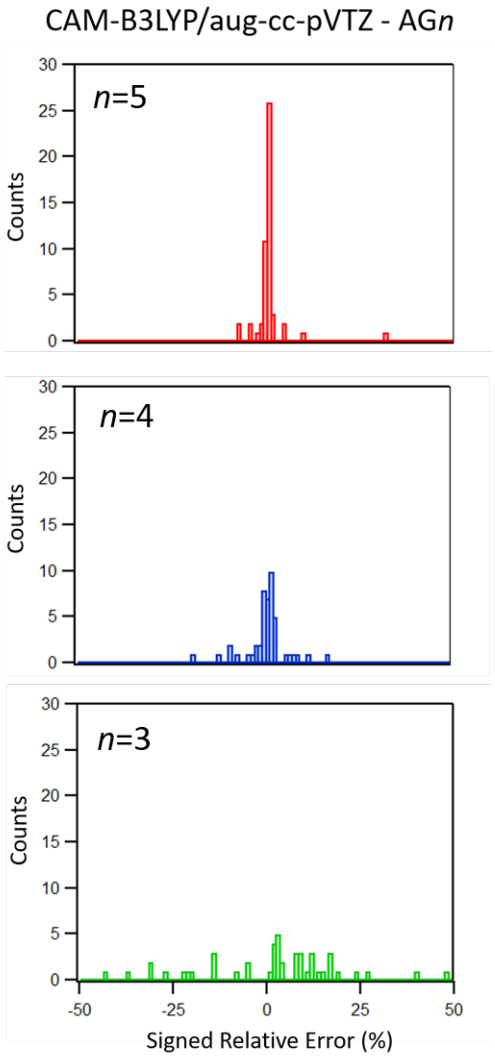


Figure 5.4: Histogram of the signed relative errors for each of the 51 molecules using the AG*n* selection criterion and the CAM-B3LYP/aug-cc-pVTZ method.

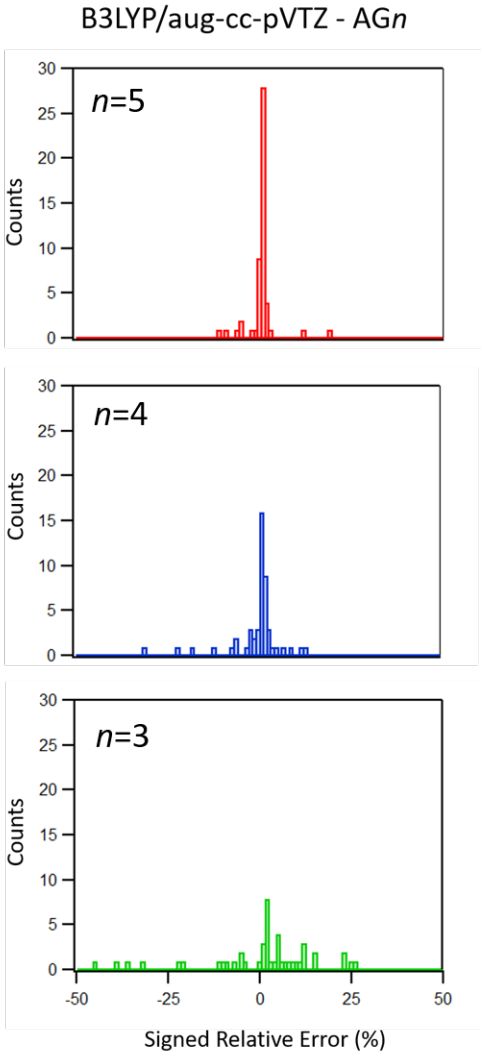


Figure 5.5: Histogram of the signed relative errors for each of the 51 molecules using the AG*n* selection criterion and the B3LYP/aug-cc-pVTZ method.

Table 5.2: Relative (%) error: MSE, MUE, Max, and σ_{MSE}, average fraction (%) of selected MOs ($\bar{N}_{sel}$) and the corresponding standard deviation ($\sigma_{\bar{N}_{sel}}$), and estimate of the average speed-up (S.U., see Eq. 5.9) for calculations with the aug-cc-pVTZ basis set.

	CAM-B3LYP			B3LYP		
n	5	4	3	5	4	3
			AGn			
MSE	0.6	4.6	-72	0.1	-2.2	-5.4
MUE	1.9	12	95	1.8	5.0	22
Max	31	230	3698	19	60	199
σ_{MSE}	5.1	38	519	4.1	11	46
$\bar{N}_{sel}$	39	23	12	39	23	13
$\sigma_{\bar{N}_{sel}}$	3	3	2	3	3	2
S.U.	17	85	600	17	79	500
			RGn			
MSE	0.3	1.0	2.7	0.0	-0.9	-11
MUE	0.8	3.7	16	1.1	3.7	23
Max	7.6	74	258	13	85	354
σ_{MSE}	1.6	12	44	2.6	13	57
$\bar{N}_{sel}$	51	34	19	40	26	14
$\sigma_{\bar{N}_{sel}}$	11	9	6	11	10	6
S.U.	7.7	26	140	15	55	380
			ASn			
MSE	0.2	1.5	-2.2	-0.1	-2.1	-6.5
MUE	0.6	6.7	33	0.6	4.6	18
Max	3.8	121	633	4.3	44	213
σ_{MSE}	1.1	20	99	1.1	11	41
$\bar{N}_{sel}$	52	30	14	52	30	14
$\sigma_{\bar{N}_{sel}}$	8	4	2	8	4	2
S.U.	7.2	37	370	7.1	37	350
			RSn			
MSE	0.3	0.2	0.9	0.0	0.0	-0.7
MUE	0.5	0.5	2.7	0.4	0.7	3.3
Max	6.9	3.2	44	3.8	5.2	32
σ_{MSE}	1.2	0.9	6.9	0.8	1.3	6.5
$\bar{N}_{sel}$	73	53	33	68	47	28
$\sigma_{\bar{N}_{sel}}$	13	12	9	16	13	9
S.U.	2.6	6.8	29	3.2	9.5	46

provides smaller spread (both in terms of σ_{MSE} and Max) and a larger speed-up than with $n = 4$ for aug-cc-pVDZ, indicating that a tighter threshold can be used for the larger basis set without sacrificing accuracy and computational gains. In fact, only the AG5 set has one molecule with error greater than 10%, molecule **6**, with an $[\alpha]_D$ of 1.3 deg dm^{-1} (g/mL)$^{-1}$ and an error of 0.4 deg dm^{-1} (g/mL)$^{-1}$. The same trends across selection criteria are observed with both basis sets, and $n = 4$ or even 3 are reasonable choices for the RSn criterion and aug-cc-pVTZ. Similar consideration apply to B3LYP/aug-cc-pVTZ, where the tight $n = 5$ choice provides both accurate results and large estimated speed-up for all selection criteria, except for RSn , where smaller n values are also acceptable. For B3LYP, there are a few cases of errors above 10%, all related to small $[\alpha]_D$ values: three molecules with AG5, and one with RG5. The three compounds for AG5 are **6**, **18**, and **34**, where the errors are -12%, 19%, and 12%, respectively, from $[\alpha]_D$ = -4, 1.5, and -9 deg dm^{-1} (g/mL)$^{-1}$, respectively; for RG5, molecule **34** has a 13% relative error corresponding to an absolute error of 1.26 deg dm^{-1} (g/mL)$^{-1}$.

5.4 Discussion and Conclusions

In this work, we show that accurate calculations of $[\alpha]_\omega$ can be performed with a subset of MOs in the solution of the CPKS equations, which determine the perturbed density used to evaluate the G' tensor. Selecting only the MOs for the excited configurations that contribute the most to the final value of the property can significantly reduce the computational cost of these calculations with minimal loss of accuracy. We propose two selection criteria for the relevant MOs, based on the guess density used in the iterative solution of the CPKS equations. With these criteria, we also propose two definitions of the selection threshold, an absolute value and a value relative to the largest element in the guess, and various numerical values for the threshold. We compiled a test set of 51 chiral organic molecules with $[\alpha]_D$ values spanning 5 orders of magnitude, and we tested the various combinations of criteria and thresholds with two functionals, CAM-B3LYP and B3LYP, and two basis sets, aug-cc-pVDZ and aug-cc-pVTZ.

We find that all of the selection criteria work fairly well, as long as the threshold value n is large

enough. The latter is the most important factor, as it correlates with the number of discarded MOs. We find that n = 5 is a very effective choice of threshold in combination with the aug-cc-pVDZ basis set. These calculations are highly accurate, with low average errors (the largest MUE across the set is 0.7% and the largest MSE is -0.2%), small σ_{MSE} (the largest is 4.1%), very small errors overall (the maximum error across all 8 sets is 9.4%), and good speed-up ranging from $2-4$x. The n = 4 threshold provides much larger speed-up, up to 11x faster than the full basis set. These calculations come with a somewhat worse accuracy than their n = 5 counterparts, as the largest MUE and MSE increase to 3.9% and -1.1%, respectively, and σ_{MSE} increases up to 8.5%. This threshold choice may be ideal for screening a large number of compounds, as the calculations can be performed quickly at a fairly high accuracy. The few large relative Max errors for the $n = 4$ calculations (see Table 5.1) are all due to small absolute values of $[\alpha]_D$. Furthermore, the corresponding absolute errors are at most 4 times smaller than the expected average error for CAM-B3LYP/aug-cc-pVDZ, and 3 times smaller than what is expected for B3LYP/aug-cc-pVDZ.[140,154] On the other hand, the $n = 3$ threshold is too loose, and it leads to low accuracy. With the larger basis set, aug-cc-pVTZ, the best threshold is $n = 5$, because the size of the basis set leads to a larger number of small density elements whose sum becomes important. Nonetheless, the accuracy of the $n = 5$ calculations with this basis set is better than the $n = 4$ equivalents with aug-cc-pVDZ, and the estimated speed-up is larger (7-8x). At a given n , our results indicate that a relative threshold is more accurate than an absolute one, and the G^2 criterion is more accurate than G^1. Thus, we recommend either the RG5 or AS5 criteria as the best compromise between computational cost and accuracy.

This proof-of-concept work shows that the selection of orbitals prior to the solution of the CPKS equations, with careful choice of the selection criterion, can lead to large estimated speed-up in calculations while accurately reproducing $[\alpha]_\omega$. These promising results now require an efficient implementation of the procedure for actual timing information. More importantly, we plan to extend this approach to coupled cluster methods, whose steeper scaling with system size compared to DFT may make the selection criteria even more beneficial for the fast evaluation of

this molecular property.

5.5 Acknowledgment

We are grateful for the support from the National Science Foundation under grant CHE-1650942.

Chapter 6

Small Basis Sets Optimized for Calculations of Optical Rotation

6.1 Introduction

Optical activity has been a topic of continuous interest due to the fundamental role chiral molecules play in pharmaceutical research. Because enantiomers of chiral molecules interact differently with a chiral environment, the correct determination of the absolute configuration (AC) of a molecule is of paramount importance in synthetic and biological fields. Due to the direct link between the structure of the molecule and its optical response, chiroptical spectroscopy and electronic structure calculations have become essential to determine the AC of chiral compounds.[2] However, the use of electronic structure methods is made difficult due to the need for highly accurate quantum chemical methods and large basis sets close to the complete basis set limit.[2] Great progress has been made in developing quantum chemical methods for performing calculations of chiroptical properties, both in density functional theory (DFT) and coupled cluster (CC) methods.[5–20] Despite these advancements, chiroptical calculations of large systems of interest are prohibitively expensive, resulting in the need for compromise between accuracy and cost of calculations.[140]

There have been few attempts to address the need for large basis sets for accurate calculations. Crawford *et al.* has implemented a localized orbital scheme in combination with the local correlation idea of Pulay and Saebø[148,149] to create cutoffs for neglecting components of the wavefunction.[64,65,155] They then tested this method with three different localization schemes, finding that while overall this method works well for pseudo-linear systems like 1-fluroalkanes, although when the dimensionality of their system increases to the larger caged structures they test, the truncation threshold needed to be reduced to maintain the same level of savings. To our knowledge, only the

the optical rotatory prediction (ORP) basis set has been specifically designed to predict chiroptical properties.[62,63] Their goal was to develop a small basis set that could be applied to a variety of optical rotation (OR) calculations of larger systems. They tested their basis set against a variety of chiral molecules, including particularly difficult flexible biological molecules.[63] Their resulting basis set was between aug-cc-pVDZ and aug-cc-pVTZ in size, and yet consistently produced values similar to aug-cc-pVTZ.[62] Howard *et al.* recently tested the ORP basis set in combination with coupled cluster calculations, and also found that they were able to reproduce aug-cc-pVTZ level results.[156]

This work builds off the previous work of Wiberg *et al.*, who constructed un-optimized small basis sets to calculate optical rotation.[61] They performed DFT and CC calculations by using the diffuse functions from aug-cc-pVDZ and aug-cc-pVTZ in combination with the 3-21G and STO-3G basis sets. Their results indicate that the core and valence functions of larger basis sets are not nearly as important as the diffuse functions. Un-augmented, the STO-3G and 3-21G basis sets failed to reproduce OR, with an average unsigned error of more than 50% for the DFT calculations. The addition of just the diffuse functions from either aug-cc-pVDZ or aug-cc-pVTZ reduced this error to just over 14%.

In this work, we have optimized the exponents of the *s*, *p*, and *d* diffuse functions from the aug-cc-pVDZ and aug-cc-pVTZ basis sets in combination with the 3-21G basis set and CAM-B3LYP. We used a training set of 21 chiral molecules with larger OR, and targeted specifically the common elements hydrogen, carbon, oxygen, and nitrogen. We then test our basis set by performing calculations using the training set at 450 nm, 633 nm, and then with B3LYP, and then perform the same set of calculations with a control set of molecules, most of which have a small OR to provide a more difficult test case.

The rest of this work is organized as follows. In section 6.2, we discuss the procedure for the optimization as well as the resulting exponents, section 6.3 contains the results of the calculations with our basis sets, and section 6.4 contains our discussion of these tests and concluding remarks.

6.2 Methods

The specific rotation of a molecule in an isotropic medium in deg dm^{-1} $(g/mL)^{-1}$ is calculated as:[93]

$$[\alpha]_\omega = \frac{(72.0\times 10^6)\hbar^2 N_A \omega}{c^2 m_e^2 M} \times \left[\frac{1}{3} Tr(G'(\omega))\right] \tag{6.1}$$

where ω is the frequency of incident light in atomic units, M is the molecular mass in amu, m_e is the mass of the electron in kg, c is the speed of light in m/s, and G' is the Rosenfeld mixed electric dipole-magnetic dipole polarizability tensor in atomic units computed with standard linear response techniques.[91,137–139]

The two compact basis sets are built by augmenting the 3-21G basis set with the diffuse functions of aug-cc-pVDZ or aug-cc-pVTZ. These basis sets are named augD-3-21G and augT3-3-21G, respectively, where T3 indicates that only three functions are included for the heavier elements (i.e., the f diffuse functions were not included, as well as the d diffuse functions for H). We left the 3-21G functions unchanged, and optimized the exponents of the diffuse functions of hydrogen, carbon, oxygen, and nitrogen, as they are very common elements in chiral organic molecules. The optimization is based on the minimization the root mean square (RMS) of the relative error in the calculation of $[\alpha]_D$ with the compact basis set, using the $\tilde{S}_k$ values computed with the corresponding correlation consistent basis set as reference, for a training set of 21 organic molecules. The minimization is based on the Newton-Raphson procedure. As an analytical derivative of $[\alpha]_D$ with respect to the exponents of the diffuse functions is complicated, we employ finite difference methods to compute the first and second derivatives of $[\alpha]_D$ (f in the following equations):

$$\begin{aligned} f'(x) &= \lim_{\delta\to 0} \frac{f(x+\delta) - f(x-\delta)}{2\delta} \\ f''(x) &= \lim_{\delta\to 0} \frac{f(x+\delta) + f(x-\delta) - 2f(x)}{\delta^2} \end{aligned} \tag{6.2}$$

The training set is formed by molecules with $[\alpha]_D$ greater than 80 deg dm^{-1} $(g/mL)^{-1}$, shown in Figure 6.1, as we used the relative error as metric for the optimization. Because our initial

tests showed that the simultaneous optimization of all basis functions was unstable, we opted for an iterative procedure where the functions for each element were optimized separately, while the others were kept fixed at the current value. We started from H, followed by C, O, and N, then the procedure was repeated for a new pass. Each element was considered converged when the change in the RMS between two successive Newton-Raphson steps was less than 5%. After 5 passes across all elements, the RMS of the error stopped decreasing, and we performed a full optimization of all 11 exponents.

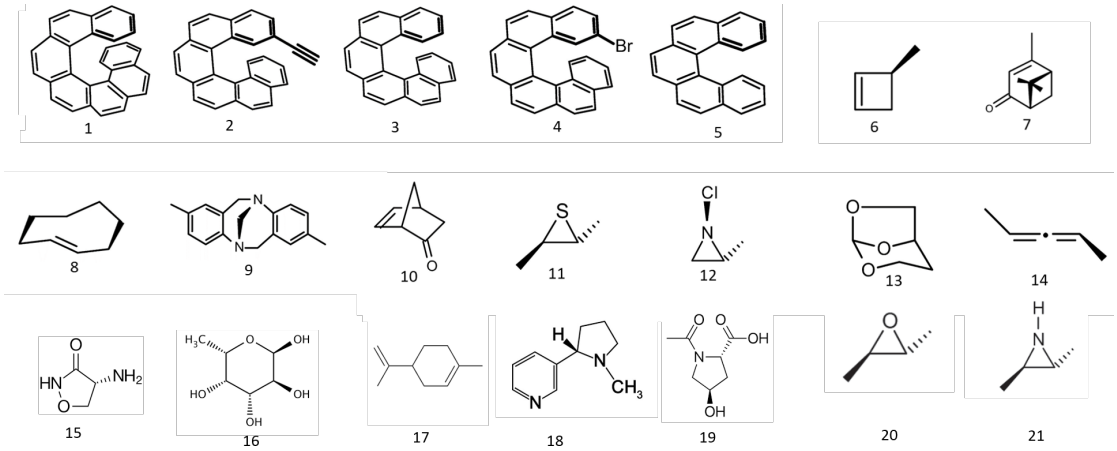


Figure 6.1: Structures of molecules in the training set.

Table 6.1: Exponents for augD-3-21G before and after optimization.

Element	H	C	O	N
		Optimized		
s	1.823999E-02	1.769533E-01	7.702713E-02	3.172590E-02
p	1.019984E-01	1.924718E-02	4.945624E-02	9.601104E-02
d		1.984507E-01	1.988384E-01	4.998359E-01
		Original		
s	2.526000E-02	4.402000E-02	7.376000E-02	5.760000E-02
p	1.020000E-01	3.569000E-02	5.974000E-02	4.910000E-02
d		1.000000E-01	2.140000E-01	1.510000E-01

In Table 6.1, we report the original and optimized exponents for augD-3-21G (a typical input for GAUSSIAN using these exponents can be found in Table 1 of the Appendix). All *s* functions except for carbon become more diffuse, while all *p* functions except for nitrogen become more diffuse. Carbon and nitrogen *d* functions become less diffuse with optimization. Both all hydrogen

Table 6.2: Exponents for augT3-3-21G before and after optimization.

Element	H	C	O	N
		Optimized		
s	2.341077E-02	1.964813E-01	3.568807E-02	3.078583E-02
p	1.016736E-01	4.837758E-02	7.727372E-02	1.544036E-01
d		1.720940E-01	2.533549E-01	5.596826E-01
		Original		
s	2.974000E-02	4.690000E-02	7.896000E-02	6.124000E-02
p	1.410000E-01	4.041000E-02	6.856000E-02	5.611000E-02
d		1.510000E-01	3.320000E-01	2.300000E-01

and all oxygen functions become more diffuse. We report the original and optimized exponents for aug-T3-3-21G in Table 6.2 (a typical input for GAUSSIAN can be found in Table 2 of the Appendix). All s functions except for carbon become more diffuse. All p functions except hydrogen become less diffuse. All d functions become less diffuse with optimization. The only difference in how the exponents change after the optimization is the carbon p, which becomes more diffuse during the augD-3-21G optimization and less so in the augT3-3-21G optimization, and in the oxygen p and d functions, which both become more diffuse during the augD-3-21G optimization and less diffuse in the augT3-3-21G optimization.

Table 6.3: Number of contracted basis functions for each element.

Element	3-21G	aug(D/T3)-3-21G	aug-cc-pVDZ	aug-cc-pVTZ
H	2	6	9	23
C	9	19	23	46
O	9	19	23	46
N	9	19	23	46

In Table 6.3, we compare the number of basis functions for each element in each basis set to give an idea of the relative cost of the calculations with each basis set. Although DFT formally scales as $O(N^4)$, practical implementations scale as $O(N^3)$ with respect to basis set size, thus, we can estimate the speedup of the calculation passing from the full correlation consistent basis set to the corresponding augX-3-21G one (X = D or T3):

$$SpeedUp = \left(\frac{N^{Target}}{N^{Opt}} \right)^3 \tag{6.3}$$

where N^{Target} is the number of basis functions of the full Dunning basis set, and N^{Opt} is the number of basis functions in our optimized basis set. We compare this estimated speedup to the actual average speedup (i.e., based on wallclock timings) for molecules **1**−**5**, which are the five largest molecules in the training set, as shorter calculation times on small molecules are more prone to variations due to computational noise.

All calculations were performed with the GAUSSIAN 09 suite of programs on the Comet machines of the XSEDE supercomputing cluster.[114] Newton-Raphson steps were performed using a locally modified version of the GAUSSIAN gauopt utility. Geometries were optimized using the CAM-B3LYP/aug-cc-pVDZ model chemistry.[115,116] Calculations of specific rotation were performed at the sodium-*D* line ($\omega = 589.3$ nm), 450 nm, and 633 nm with CAM-B3LYP[117–119] and at the sodium-*D* line with B3LYP, using the aug-cc-pVDZ, aug-cc-pVTZ, and the optimzied and un-optimized versions of the augX-3-21G basis sets.

6.3 Test Calculations

We performed a series of tests to determine the transferability of the augD-3-21G and augT3-3-21G basis sets. These tests include calculations on the molecules in the training set with CAM-B3LYP at two other wavelengths, 450 and 633 nm, and with B3LYP at 589.3 nm. In order to have a more comprehensive test of our basis set, we chose a second set of molecules, labeled the control set (see Figure 6.2). These molecules have $[\alpha]_D$ mostly below 100 deg dm^{-1} (g/mL)$^{-1}$, providing a more stringent test for the new basis sets. All $[\alpha]_\lambda$ values are reported in Tables 3 and 6 of the Appendix.

As mentioned in the previous section, the $[\alpha]_\lambda$ values computed with the full aug-cc-pVDZ are used as reference for the augD-3-21G basis set, and those with aug-cc-pVTZ are used as reference for the augT3-3-21G basis set. The performance of the optimized exponents in the reduced basis set is compared to those obtained by using the original values of the exponents. Results for the training set are reported in terms of relative (%) errors, whereas control set values are presented as absolute errors (calculated as $[\alpha]_\lambda^{Opt} - [\alpha]_\lambda^{Ref}$), because the $[\alpha]_\lambda$ magnitude is small. Tables

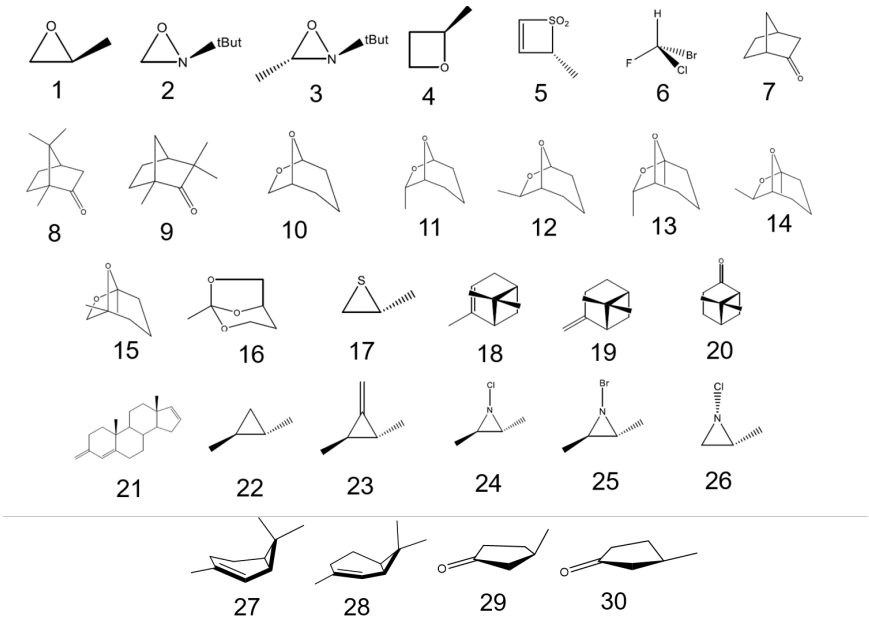


Figure 6.2: Structures of molecules in the control set.

collect the mean signed error (MSE), mean unsigned error (MUE), maximum error (Max), and the signed error standard deviation (σ). We also evaluate the relative computational cost of the reduced basis sets compared to the full Dunning ones using a theoretical estimate of the speed-up as in Eq. 6.3, and wallclock relative timings for molecules **1**-**5** of the training set. These calculations were performed on the same machine with the same number of processors and memory.

It is important to discuss the errors of the reduced basis set in the context of the expected error of the full basis sets. The latter is known for aug-cc-pVDZ thanks to Srebro *et al.*[140] and Stephens *et al.*[154] Srebro *et al.* benchmarked optically active compounds with B3LYP/aug-cc-pVDZ and CAM-B3LYP/aug-cc-pVDZ against experimental results using a test set of molecules with $[\alpha]_D$ ranging from -400 to +500 deg dm^{-1} (g/mL)$^{-1}$.[140] They found average deviations of 25-30 deg dm^{-1} (g/mL)$^{-1}$ with CAM-B3LYP, and 20-25 deg dm^{-1} (g/mL)$^{-1}$ for B3LYP, also in agreement with previous findings from Stephens *et al.*.[154]

Table 6.4: Relative error (% MSE, MUE, Max, and σ) for the training set with the augD-3-21g basis set.

	CAM-B3LYP			B3LYP
λ	589.3 nm	450 nm	633 nm	589.3 nm
Optimized				
MSE	-1.0	-1.9	-0.8	-1.9
MUE	3.7	4.5	3.7	3.8
Max	7.3	11.1	7.2	8.7
σ	4.4	5.4	4.3	4.5
Original				
MSE	-0.7	-1.3	-0.6	-1.2
MUE	6.8	7.8	6.7	6.7
Max	18.0	23.6	17.4	19.4
σ	8.7	10.1	8.6	8.7

6.3.1 augD-3-21G

A statistical analysis of the results for the augD-3-21G basis set with the training set is presented in Table 6.4. The results obtained with CAM-B3LYP at $\lambda = 589.3$ nm (first column in the table), i.e., the optimization set, show a reduction of the error by about half with respect to those obtained with the original exponents for all statistical metrics. For instance, MUE and σ are reduced from 6.8% to 3.7% and from 8.7% to 4.4%, respectively. More interestingly, the control calculations at the other wavelengths and with B3LYP follow the same trends, with the MUE and σ also decreasing by a factor of about 2, from 7.8% and 10.1% to 4.5% and 5.4% at 450 nm, from 6.7% and 8.6% to 3.7% and 4.3% at 633 nm, and from 6.7% and 8.7% to 3.8% and 4.5% when using B3LYP at 589.3 nm, respectively. All calculations of the training set with the reduced basis sets, both before and after optimization, produce $[\alpha]_\lambda$ with the same sign as aug-cc-pVDZ.

The statistical data for the absolute errors using the control set are reported in Table 6.5. The results with CAM-B3LYP at 589.3 nm do not present a large change in the error after the optimization, as both the MUE and σ change by less than 1 deg dm^{-1} (g/mL)$^{-1}$. This is because the results obtained with the original exponents are already extremely good for this series of molecules. The calculations at the other two wavelengths and with B3LYP provide similar trends, with the results obtained with the optimized exponents being only slightly worse (generally less than 1 deg dm^{-1}

Table 6.5: Absolute error (MSE, MUE, Max, and σ) for the control set with the augD-3-21g basis set.

	CAM-B3LYP			B3LYP
λ	450 nm	589.3 nm	633 nm	589.3 nm
		Optimized		
MSE	1.1	1.2	0.3	5.5
MUE	10.5	4.8	4.5	10.8
Max	28.1	17.2	13.2	58.9
σ	13.6	6.6	6.2	18.4
		Original		
MSE	1.7	0.4	1.0	6.1
MUE	9.7	5.5	3.7	9.3
Max	32.4	15.4	15.1	53.4
σ	12.7	7.2	5.7	15.1

$(\text{g/mL})^{-1}$) compared to those with the original exponents. Only one calculation using the control set across all sets of calculations produces an incorrect sign of $[\alpha]_\lambda$ relative to the aug-cc-pVDZ results. This is for molecule **20** at 450 nm, which has $[\alpha]_{450} = -1.85$ deg dm^{-1} $(\text{g/mL})^{-1}$ with the full aug-cc-pVDZ basis set. The augD-3-21G results gives $[\alpha]_{450} = 2.08$ deg dm^{-1} $(\text{g/mL})^{-1}$ with the original exponents and 11 deg dm^{-1} $(\text{g/mL})^{-1}$ with the optimized exponents. Both values are well within the average error expected for this level of theory, see Section 6.3. At 450 nm, the Max error with the optimized basis set in Table 6.5 is due to molecule **22**, which has $[\alpha]_{450} = 98.13$ deg dm^{-1} $(\text{g/mL})^{-1}$ with aug-cc-pVDZ, while with the reduced basis set $[\alpha]_{450} = 87.13$ before the optimization and $= 70.04$ deg dm^{-1} $(\text{g/mL})^{-1}$ after optimization. The Max value with the original exponents is due to molecule **3**, which at this wavelength has $[\alpha]_{450} = -153.64$ deg dm^{-1} $(\text{g/mL})^{-1}$ with the full basis set, and errors of 32.4 with the original exponents and 22.1 deg dm^{-1} $(\text{g/mL})^{-1}$ with the optimized exponents of the reduced basis set. The two large Max errors for B3LYP in Table 6.4 are due to two different molecules, molecule **30** for the original exponents ($[\alpha]_D = -274.59$, -327.97, and -329.6 deg dm^{-1} $(\text{g/mL})^{-1}$ with aug-cc-pVDZ, augD-3-21g with the original, and optimized exponents, respectively), and molecule **29** for the optimized exponents ($[\alpha]_D = 235.89$ with aug-cc-pVDZ, and errors of 53.4 and 55.0 deg dm^{-1} $(\text{g/mL})^{-1}$ for augD-3-21g with the original and the optimized exponents, respectively). Note that the Max errors are due

to the molecules with larger values of specific rotation, so that their relative errors are similar to those reported in Table 6.4 for the molecules in the training set, which also have larger values of $[\alpha]_\lambda$.

In Table 7 of the Appendix, we present the number of basis functions for each molecule with both aug-cc-pVDZ and augD-3-21G. Using Eq. 6.3, we estimate that a calculation with augD-3-21G should be on average 2 times faster than an aug-cc-pVDZ calculation. Calculations of the 5 largest molecules in our set are 3-5 times faster with the smaller basis set (4.25x speedup on average).

6.3.2 augT-3-21G

In this section, we discuss the performance of augT3-3-21G compared to the full aug-cc-pVTZ basis set. We also present the results with the reduced basis set and the original exponents, with a variant of the reduced basis set that includes all diffuse functions with original exponetns (augT4-3-21G), and with the full aug-cc-pVDZ results. The latter provides an important test because the augT3-3-21G basis set is significanly more compact than aug-cc-pVDZ (in fact, it has the same number of functions per element as the augD-3-21G basis set), and it can provide considerable computational savings.

The statistical results for the relative error with the training set are reported in Table 6.6, where again the first column contains the data from the actual exponent optimization. Not surprisingly, the augT3-3-21G results with CAM-B3LYP at 589.3 nm are considerably better with the optimized exponents than those with the original exponents across the board. More interestingly, the optimized augT3-3-21G basis set outperforms augT4-3-21G as well as the full aug-cc-pVDZ on virtually all metrics. This excellent performance is also replicated at the other two wavelengths with CAM-B3LYP, and at 589.3 nm with B3LYP. Remarkably, the calculations with the full aug-cc-pVDZ basis set are only marginally better than those with augT3-3-21G for the MUE metric, but they are significaly worse for all other metrics. All calculations with the small basis sets produce $[\alpha]_D$ values of the same sign as the aug-cc-pVTZ basis set.

Table 6.6: Relative error (% MSE, MUE, Max, and σ) for the training set with the augT3-3-21g basis set.

	CAM-B3LYP			B3LYP
λ	589.3 nm	450 nm	633 nm	589.3 nm
Optimized augT3-3-21G				
MSE	-1.6	-2.2	-1.4	-1.4
MUE	4.4	4.9	4.3	4.2
Max	10.1	12.7	8.8	8.3
σ	5.3	6	5	4.9
Original augT4-3-21G				
MSE	1.1	1	1.3	1.3
MUE	5.4	5.7	5.1	5
Max	17.8	18.6	15	14.5
σ	6.6	7.1	5.9	5.8
Original augT3-3-21G				
MSE	1.9	0.7	1.7	1.8
MUE	12.2	14	12.3	12.1
Max	48.8	52	46.2	45.4
σ	17.7	19.6	17.3	17
aug-cc-pVDZ				
MSE	3.1	3.3	3.1	3
MUE	4.2	4.4	4	4
Max	22	22.4	20.9	20.7
σ	6.9	7.3	6.6	6.6

Table 6.7 shows the statistics for the absolute error of the control set. The optimization clearly improves the performance of the augT3-3-21G basis set compared to the original exponents, with a decrease of the MUE by more than 30%. A performance improvement is also recorded compared to the augT4-3-21G basis set with original exponents across all metrics. However, the most interesting comparison is again with aug-cc-pVDZ. The results in Table 6.7 show that augT3-3-21G is definitely competitive with the considerable larger aug-cc-pVDZ basis set, with error values that are a few deg dm^{-1} (g/mL)$^{-1}$ larger or smaller than aug-cc-pVDZ. The augT3-3-21G basis set provides the wrong sign of $[\alpha]_{450}$ compared to aug-cc-pVTZ for two molecules, **20** and **25**, because the reference results are very close to 0 (molecule **20** has $[\alpha]_{450}$ = -3.24, 3.6, -0.65, 15.2, and -1.85 deg dm^{-1} (g/mL)$^{-1}$ with aug-cc-pVTZ, augT4-3-21G, unoptimized augT3-3-21G, optimized augT3-3-21G, and aug-cc-pVDZ, respectively, while molecule **25** has $[\alpha]_{450}$ = -2.21, 38.51,

Table 6.7: Absolute error (MSE, MUE, Max, and σ) for the control set with the augT3-3-21g basis set.

	CAM-B3LYP			B3LYP
λ	450 nm	589.3 nm	633 nm	589.3 nm
Optimized augT3-3-21G				
MSE	1.4	1	0.9	1
MUE	10.4	5.9	5	6.4
Max	27.5	15.1	13	17
σ	13.8	7.6	6.5	8.3
Original augT4-3-21G				
MSE	6.4	3.1	2.6	3.3
MUE	12.4	6.4	5.4	6.4
Max	51.2	27	23	29
σ	17.5	8.9	7.6	9.2
Original augT3-3-21G				
MSE	7.6	3.7	3.1	3.6
MUE	18.4	9.2	7.8	9.6
Max	65.7	35.9	30.8	37.9
σ	26.6	13.6	11.5	14.3
aug-cc-pVDZ				
MSE	3.2	1.7	1.4	1.4
MUE	5.9	3.1	2.6	8.2
Max	47.5	25.3	21.5	50.3
σ	10.6	5.7	4.8	15.8

63.15, 2.39, and -0.6 deg dm^{-1} (g/mL)$^{-1}$ with aug-cc-pVTZ, augT4-3-21G, unoptimized augT3-3-21G, optimized augT3-3-21G, and aug-cc-pVDZ, respectively). At 589.3 nm with B3LYP, the calculation for molecule **23** has the wrong sign compared to the reference with all basis sets, including aug-cc-pVDZ. Again, this is due to the small magnitude of $[\alpha]_D$: -3.73, 25.22, 34.13, 9.99, and 11.77 deg dm^{-1} (g/mL)$^{-1}$ with aug-cc-pVTZ, augT4-3-21G, unoptimized augT3-3-21G, optimized augT3-3-21G, and aug-cc-pVDZ, respectively. All calculations with CAM-B3LYP using the optimized augT3-3-21G basis set produce the same sign of $[\alpha]_\lambda$ as the aug-cc-pVTZ basis set at 589.3 nm and 633 nm. The largest Max error with augT3-3-21G with optimized exponents is 27.5 deg dm^{-1} (g/mL)$^{-1}$ at 450 nm, see Table 6.7, which is due to molecule **28**; this molecule has a large $[\alpha]_{450}$ of 403.74, and the errors with the other basis sets are 2.9 with augT4-3-21G, 17.1 with augT3-3-21G (original exponents), and -1.2 deg dm^{-1} (g/mL)$^{-1}$ with aug-cc-pVDZ.

However, note that the Max errors are almost always larger with the other basis sets, including aug-cc-pVDZ.

Table 8 of the appendix presents the total number of basis functions for each molecule with both the aug-cc-pVTZ and augT3-3-21G basis sets. From these values, we expect that an augT3-3-21G calculation should be on average 22 times faster than an aug-cc-pVTZ calculation. This agrees with the relative times of our calculations for the 5 largest molecules, with speed-ups between 18x and 26x.

6.4 Discussion and Conclusions

In this work, we propose two compact basis sets obtained from combining the standard 3-21G basis with diffuse functions obtained from the aug-cc-pVDZ basis set (augD-3-21G), and from the aug-cc-pVTZ basis set (augT3-3-21G). For the latter, we exclude the diffuse functions of highest angular momentum, so that augD-3-21G and augT3-3-21G include the same number of functions per element. The exponents of the diffuse functions for four elements, H, C, O, and N, are optimized to minimize the RMS difference from the $[\alpha]_D$ values computed with CAM-B3LYP and the corresponding full Dunning basis sets, using a training set of 21 organic molecules. The new basis sets are then tested on the same training set but at different wavelengths, 450 nm and 633 nm, with B3LYP at 589.3 nm, and on a control set of different molecules with all the previous functionals and wavelentghs. The results are compared with those obtained with the reduced basis set but using the original exponents, and in the case of augT3-3-21G, also against the full aug-cc-pVDZ.

The tests using the training set show that the augD-3-21G basis set with optimized exponents is improved compared to using the original exponents, as the MSE and σ are reduced by nearly half at all frequencies and with all functionals. The tests with the control set do not show a particular improvement with the optimization of the exponents, because the results with the original values are already rather close to those with aug-cc-pVDZ (MUE and σ for the optimized and original exponents are almost all within 1 deg dm^{-1} (g/mL)$^{-1}$ of each other). The difference between the augD-3-21G and aug-cc-pVDZ results are also on average smaller than the accuracy of the full

basis set compared to experiment with both functionals. In fact, the average expected error relative to experimental values from B3LYP is 20-25 deg dm^{-1} $(g/mL)^{-1}$ and the average expected error from CAM-B3LYP is 25-30 deg dm^{-1} $(g/mL)^{-1}$,[140,154] whereas the average errors with augD-3-21G with respect to aug-cc-pVDZ are 8.6 and 4.9 deg dm^{-1} $(g/mL)^{-1}$. From the relative sizes of the basis sets, and using the $O(N^3)$ scaling of DFT, the augD-3-21G basis set should be at least 2 times faster than the aug-cc-pVDZ basis set (actual timings with the five largest molecules in the set showed average speed-up of 4x).

The optimization of the exponents for the augT3-3-21G basis set leads to a large improvement compared to the results with the original exponents for both the training and the control sets. The optimized augT3-3-21G basis performs better than the augT4-3-21G basis set, which includes all diffuse functions from aug-cc-pVTZ with original exponents. More importantly, augT3-3-21G provides results that are comparable to (control set, Table 6.7) or better than (training set, Table 6.6) the full aug-cc-pVDZ compared to aug-cc-pVTZ, at a fraction of the computational cost. Using the relative sizes of the basis sets for each molecule and the $O(N^3)$ scaling of DFT, the augT3-3-21G basis set is expected to be nearly 22x faster than the full aug-cc-pVTZ basis set (which is confirmed by the actual relative timings with the five largest molecules in the set).

Both the augT3-3-21G and augD-3-21G basis set optimizations change the exponents in mostly similar ways. The main differences are the *d* diffuse functions of oxygen, which are much more diffuse after the augD-3-21G optimization, but change very little after the augT3-3-21G optimization. The carbon and oxygen p functions also change differently for the two basis sets; however, the change in the value of the exponents is overall small. All *s* functions become more diffuse except for carbon, while all *p* functions, except for hydrogen and the augD-3-21G oxygen, become less diffuse. All *d* functions become less diffuse in both basis sets and for all elements.

In summary, both the augD-3-21G and augT3-3-21G basis sets perform well compared to the full Dunning basis sets, and they allow calculations of specific rotation of the same quality of larger basis sets at a fraction of the computational work. This is particularly true for augT3-3-21G, which provides results as close to aug-cc-pVTZ as the full aug-cc-pVDZ basis set. These results

are indeed quite promising, and we are now testing these reduced basis set using coupled cluster methods and against the ORP basis set.

6.5 Acknowledgment

T.A. and M.C. are grateful for the support from the National Science Foundation under grant CHE-1650942, and from the the Extreme Science and Engineering Discovery Environment under project CHE170088.

Chapter 7

Theoretical Investigations of Helicenes

(This work is partially taken with the permission of Tal Aharon, Marco Caricato, Kévin Martin, and Narcis Avarvari from *Chem. Eur. J.*, **2017,** *23*, 437-446. Supporting information is available online.[81]

All experimental work detailed in this Chapter was performed by the Avarvari group at the Universitè D'Angers, France, and their collaborators.)

7.1 Introduction

Heterohelicenes are a particularly interesting family of the nonplanar conjugated aromatic helical molecules known as helicenes,[67,157,158] as they combine the inherent properties of carbohelicenes, such as strong optical rotation and circular dichroism (CD),[20,68,102] nonlinear optical properties,[103,104] self-assembly abilities towards supramolecular materials,[69,159] with the specificity of the heterocycle or the corresponding heteroatom. A large majority of the reports on heterohelicenes concern thiahelicenes, in which thiophene units replace totally[160–164] or partially the benzene rings,[165–169] or azahelicenes containing pyridine rings,[170–174] together with their metal complexes.[175–177] However, other heterocyclic units such as carbazole,[178–180] pyrane,[181,182] xanthenium,[183] quinacridinium,[184,185] pyridinium,[186–188] phenoxazine,[189] phenothiazine,[190] phosphole,[191–193] dibenzofuran,[194] or the more exotic azaborine[75] and silole[195] have been included relatively recently in helical structures for different functions and properties. The presence of certain heterocycles in the helical skeleton provides electroactive character, with the possibility to access chiroptical redox switches, or allows modulation of the luminescence properties. Con-

cerning the former, besides thiahelicenes[161] and helquats,[187] examples of electroactive helicenes that have been studied for redox modulation of the CD signal are still rare in the literature[78,196] and include organometallic derivatives of [6]helicene,[85,197] tetrathiafulvalene-helicenes,[198] and helicene-quinones.[199] For a phenothiazine-based double hetero[4]helicene, a crystalline radical cation salt has been isolated and structurally characterized.[190] On the other hand, it is known that fluorescence quantum yields of carbohelicenes are generally very low due to efficient intersystem crossing (ISC) from singlet to triplet excited states,[200] yet the presence of heterocycles such as carbazoles[178–180] and siloles[195] that themselves have luminescent properties strongly enhances the fluorescence emission of the corresponding heterohelicenes. Moreover, when thiahelicenes are fused with electron-acceptor quinoxaline units,[166] or upon oxidation of the sulfur atoms to sulfone groups,[201] high fluorescence quantum yields and circularly polarized luminescence are observed, as a consequence of an increased energy gap between the lowest singlet and triplet states. Interestingly, the emission of an S-shaped double azahelicene is stronger than that of the simple azahelicene congener.[173] A particularly interesting heterocyclic unit for its emission properties and electron-acceptor character is benzothiadiazole (BTD),[202,203] which has been extensively used during the last decade in the design of materials for red-light emission,[204,205] organic field-effect transistors,[206–209] and photovoltaics,[210–212] or associated in donor–acceptor dyads with tunable luminescence.[212–214] However, thiadiazole units have been never fused to helicene scaffolds to the best of our knowledge. We describe herein the DFT simulations of the chiroptical and photophysical properties that helped the characterization of enantiopure thiadiazole-[7]helicene and its S-shaped double [4]helicene, together with the smaller congeners thiadiazole-[5]helicene and thiadiazole-benzanthracene, respectively, as first representatives of a new family of electron-poor heterohelicenes. These compounds were synthetized and characterized experimentally by our collaborators in France. We also investigated a series of [N]helicene-BTD-[N]helicene fused molecules, also synthesized and characterized by our collaborators in France, as they are expected to perform well as circularly polarized light emitters to create organic light emitting diodes.[215]

7.2 Computational Details

All calculations were performed with a development version of the GAUSSIAN suite of programs.[114] Geometries were optimized using the CAM-B3LYP[115]/aug-cc-pVDZ model chemistry with D3 dispersion corrections. Calculations of absorption, ECD, and emission spectra were performed at the same level of theory in the length gauge formalism using gauge-including atomic orbitals[19,94,95]. Solvent effects were accounted for using the polarizable continuum model[131–133] (PCM) and the SMD radii.[216]

7.3 Results

7.3.1 Chiroptical properties of thiadiazole-[7]helicene

Electronic CD (ECD) measurements for the two enantiomers (P)-(+)- and (M)-(-)-thiadiazole[7] helicene show the expected mirror image relationship, with a series of positive bands at 444, 370, 323 nm and two negative bands at 297 and 241 nm for (P)-thiadiazole[7]helicene, and vice versa for (M)-thiadiazole[7]helicene (Figure 7.1, top). The UV/Vis absorption spectra (see also Figures S12 and S13 of the Supporting Information) show, as in the case of the ECD spectrum, a series of high energy bands, the most intense of which is centered at 260 nm, very likely arising from a helicene-based $\pi \rightarrow \pi^*$ transition,[198] and then at 300 and 360 nm. In the lower-energy region the weak band appearing at 445 nm (e=1300 M^{-1} cm^{-1}) corresponds to an intramolecular charge transfer (ICT) transition from helicene to the thiadiazole acceptor unit (see below). The calculated ECD and absorption spectra (Figure 7.1, bottom) excellently reproduce the experimental spectra of (P)-thiadiazole[7]helicene, although peak maxima are blueshifted by about 40 nm. All bands correspond to $\pi \rightarrow \pi^*$ excitations that involve different orbitals. The 444 nm band is a HOMO $\rightarrow$ LUMO charge-transfer transition from the helicene structure to the thiadiazole unit. The 370 nm band is also a charge-transfer transition from the helicene to the thiadiazole, which can be primarily described as HOMO-2 $\rightarrow$ LUMO and HOMO-1 $\rightarrow$ LUMO+1. The band at 323 nm is primarily a HOMO $\rightarrow$ LUMO+3 transition, and the negative band at 297 is a transition from the thiadiazole

to the helicene backbone (predominantly HOMO-2 → LUMO+1 and HOMO-1 → LUMO+2).

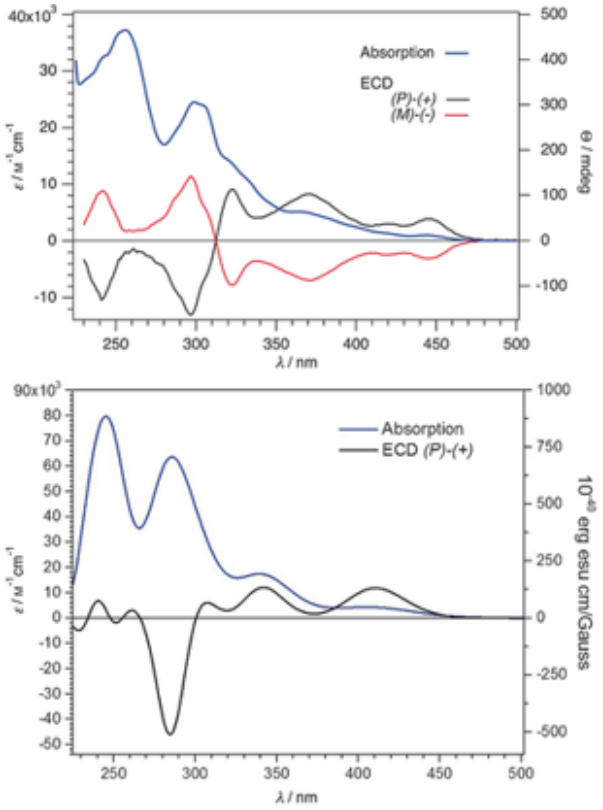


Figure 7.1: Experimental UV/Vis and ECD spectra of **1** in dichloromethane (5 x 10^{-5} M) at room temperature (top) and theoretical UV/Vis and ECD spectra of (P)-**1** (bottom); DFt/CAM-B3LYP/aug-cc-pVDZ, solvent: CH_2Cl_2, solvation model: PCM with SMD radii. Excitations appear at 410.67, 342.84, 301.5, 285.19, 262.54, and 243.2nm.

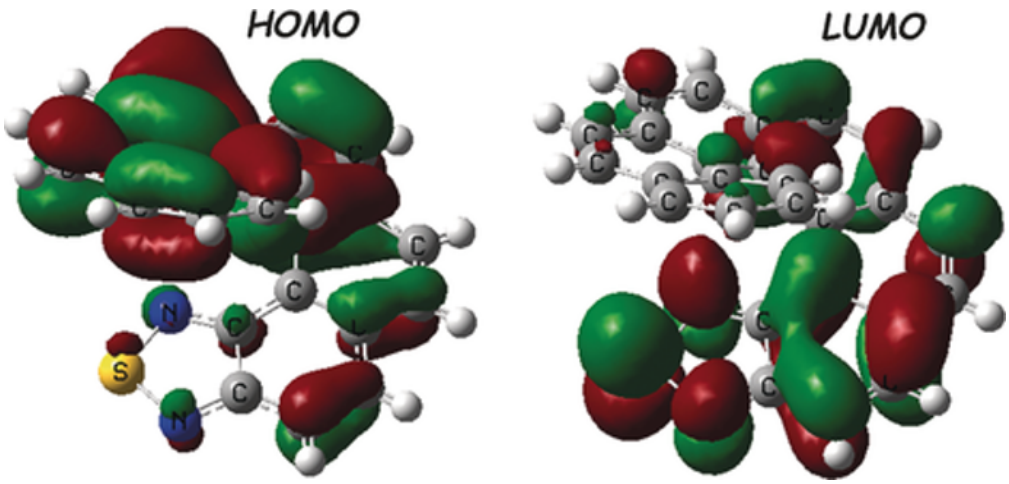


Figure 7.2: HOMO (-6.85 eV) and LUMO (-1.47 eV) orbitals for (P)-**1** (DFT/CAM-B3LYP/aug-cc-pVDZ).

Interestingly, in the present case ICT could be operative through bond, as the molecule is conjugated, or through space thanks to the intramolecular π-stacking interaction between the thiadiazole and the two terminal benzene units, as suggested by the solid-state structure (see Figure 1 in Ref. 81). This hypothesis is supported by the composition of the frontier orbitals, with the HOMO especially developed on the three terminal benzene rings, while the LUMO is concentrated on the BTD part Figure 7.2. All of the orbitals and their respective energies are reported in Figure S14 and Table S7 of the Supporting Information.

7.3.2 Photophysical properties and theoretical investigations of thiadiazole-[7]helicene, double[4]helicene, thiadiazole-[5]helicene, and thiadiazole-benzanthracene

Since the BTD unit was introduced into the structures of thiadiazole-[7]helicene (**1**), double[4]helicene (**2**), thiadiazole-[5]helicene (**3**), and thiadiazole-benzanthracene (**4**) for its luminescence, beside the electron-poor character, the Avarvari group determined the emission properties of the helical thiadiazoles. However, the luminescence of helicenes can be strongly decreased by ISC. On excitation at λ_{ex} =365 nm, both enantiomers of **1** show a weak emission band centered at 490 nm (20400 cm^{-1}), as shown in the Supporting Information (Figure S12 for (M)-**1** and Figure S13 for (P)-**1**). The corresponding excitation spectra show the main features of the absorption spectra, but they are not totally superimposable, probably due to some underlying luminescence of an impurity. Thus, the experimental quantum yield of $< 0.1\%$ must be regarded as the upper limit for the luminescence of thiadiazole-[7]helicene, which is evidently much lower than the usual values for BTD derivatives.[217] To investigate the quenching of the fluorescence, we optimized the geometry of the first excited singlet state of (P)-**1**, and the energy levels of singlet and triplet states are shown in Figure 7.3. Assuming that ISC occurs from the relaxed S1 geometry, the plot in Figure 7.3 shows a favorable energy alignment with nearby triplet states. To study the effect of spatial superposition of the terminal rings, we performed similar calculations with smaller congeners by removing the

terminal benzene ring or the thiadiazole ring. The data for these smaller systems, reported in the Supporting Information (Figures S15 and S16 and Tables S8–S10) show a worse energy alignment for ISC, in agreement with results for [5]helicene and monoaza[5]helicenes.[218]

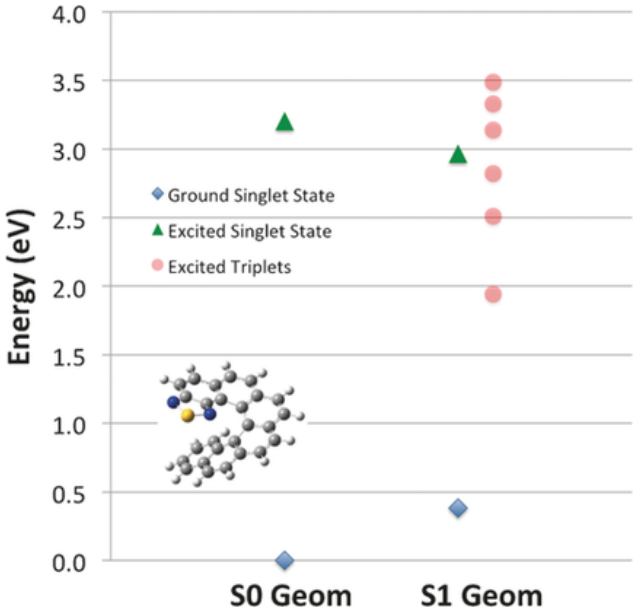


Figure 7.3: Calculated energy levels [eV] of ground and excited singlet and triplet states for (P)-**1**, evaluated at the ground-state optimized geometry (S0 Geom) and at the first singlet excited state optimized geometry (S1 Geom). The ground-state energy at S0 Geom was used as reference. The energy levels of the triplet states are slightly shifted for clarity. All excited singlet and triplet energy levels were computed with linear response methods.

In striking contrast to **1**, S-shaped double helicene **2** exhibits luminescence in dichloromethane solution both at room temperature, with a broad band centered at 528 nm (Figure 7.4, top), and a slightly structured band at 77 K (Supporting Information, Figure S17), with a relatively high emission quantum yield of 5.4% at room temperature when excited at $\lambda_{ex} = 420$ nm. The corresponding excitation spectrum is perfectly superimposable on the absorption spectrum, and this indicates that the luminescence is indeed exclusively from compound **2**.

The ground state and the lowest excited state of **2** were optimized (Supporting Information, Figure S18). The calculated absorption and emission spectra reproduce accurately the experimental ones (Figure 7.4, bottom). From the calculated spectra we were able to assign all of the transitions as primarily $\pi \rightarrow \pi^*$ charge-transfer transitions. The absorption peak at 330 nm is due

to a transition from the HOMO to the LUMO+2 on the helicene unit (see Figure S19 and Table S11 of the Supporting Information). The lower-energy charge-transfer transitions are all from the helicene to the thiadiazole unit, and differ only in the orbitals involved. The large peak at 360 nm comes primarily from HOMO-1 $\rightarrow$ LUMO transition, and the broad set of peaks around 400 nm originates from a similar charge transfer between HOMO and LUMO. For the emission spectrum, the calculated peak at λ_{ex} = 491.3 nm (corresponding to the experimental peak at λ_{em} = 528 nm) is characterized by charge transfer from the LUMO on the thiadiazole unit back to the HOMO on the helicene.

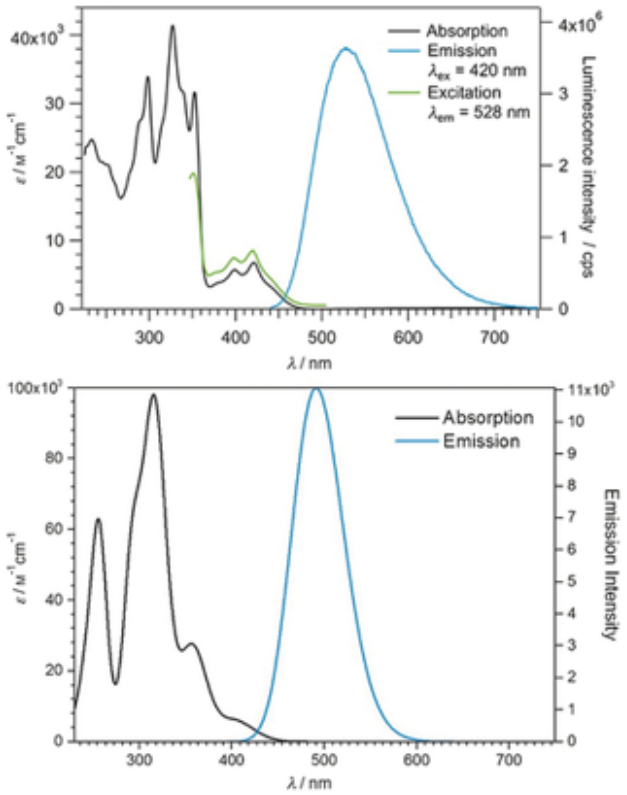


Figure 7.4: Absorption, emission, and excitation spectra of **2** (c=2.58 x 10^{-5} M in CH_2Cl_2 solution) at RT; for emission λ_{ex} = 420 nm, for excitation λ_{em} = 528 nm (top). Calculated absorption and emission spectra of **2**; DFT/ CAM-B3LYP/aug-cc-pVDZ, solvent: CH_2Cl_2, solvation model: PCM with SMD radii and nonequilibrium solvation; absorption peaks at 316.1, 357.7, and 404.7 nm for λ_{em} =491.3 nm (bottom).

As for **1**, a very weak emission band with maximum at 474 nm was recorded for **3** by excitation in solution at 386 nm (Supporting Information, Figure S20). The slight differences of

the absorption spectrum with the excitation spectrum (measured at emission wavelength 474 nm) might suggest that this luminescence is probably not real. On cooling the solution at 77 K the luminescence becomes stronger and the spectrum more structured with a slight shift to lower energy (Supporting Information, Figure S21). Likewise, the excitation spectrum does not fit perfectly the absorption spectrum; therefore, assignment of this luminescence to some impurity cannot be excluded. Nevertheless, the calculated absorption and emission spectra of **3** are in good agreement with the experimental ones (see Figures S20 and S22 of the Supporting Information for the optimized geometries of 3 in the ground state and singlet excited state), and the question of the origin of the experimentally observed luminescence remains open. The photophysical properties of compound **4** resemble those of its longer congener **2**. It exhibits luminescence in dichloromethane solution both at room temperature (Figure 7.5, top) and at 77 K (Supporting Information, Figure S24) with an emission quantum yield of 6.5%, slightly higher than that of **2**. The emission band in solution is centered at 515 nm for excitation at λ_{ex} = 396 nm. Since the excitation spectra are perfectly superimposed on the absorption spectra, this compound exhibits real luminescence, which is further supported by theoretical calculations (see Figure 7.5, bottom, and Figure S25 of the Supporting Information for the optimized geometries of **4** in the ground state and singlet excited state).

Once again, the agreement between the experimental and calculated spectra allowed us to assign the peaks in the absorption and emission spectra. Similar to the longer helicene **2**, the transitions in **4** are primarily $\pi \rightarrow \pi^*$. The peak at 406.7 nm is mainly due to charge transfer from the HOMO localized on the helicene unit to the LUMO localized on the thiadiazole group (see Supporting Information, Figure S26). The peak at 341.5 nm is a transition from the HOMO-1, which is more spread out across the entire molecule, to the LUMO. Finally, the peak at 296 nm is a combination of two transitions, both from the HOMO but one to the LUMO+1 and the other to the LUMO+2. The emission peak is a LUMO to HOMO charge transfer from the thiadiazole to the helicene unit.

Following the work in Ref. 81, the Avarvari group synthesized a series of M-[N]helicene-BTD-

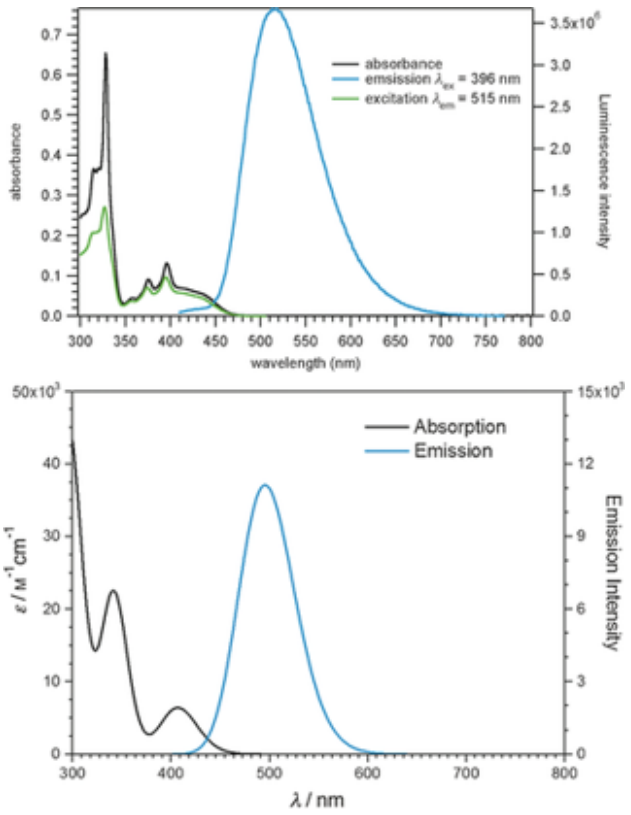


Figure 7.5: Absorption, emission, and excitation spectra of **4** (c=2.58 x 10^{-5} M in CH_2Cl_2 solution) at RT; for emission λ_{ex} = 396 nm, for excitation λ_{em} = 515 nm (top). Calculated absorption and emission spectra of **4**; DFT/ CAM-B3LYP/aug-cc-pVDZ, solvent: CH_2Cl_2, solvation model: PCM with SMD radii and nonequilibrium solvation; absorption peaks at 406.7, 341.5, and 296 nm for λ_{em} =495.2 nm (bottom).

A) [4]H-BTD-[4]H B) [5]H-BTD-[5]H C) [6]H-BTD-[6]H

Figure 7.6: A) M-[4]H-BTD-M-[4]H, B) M-[5]H-BTD-M[5]H, and C) M-[6]H-BTD-M[6]H

M-[N]helicene molecules, with N = 4, 5, 6 (reported as M-[N]H-BTD-M-[N]H in Figure 7.6). Since these are conformationally flexible molecules, multiple minima are present, which need to be accounted for to reproduce the experimental spectra. We first determined all stable conformers at CAM-B3LYP/6-31G(d) level. We then selected enough conformers such that the total Boltzmann contribution was greater than 70% (2 conformers for each molecule), and re-optimized these structures with the aug-cc-pVDZ basis set. We computed the spectra for the selected conformers with the larger basis set, with Boltzmann weighting evaluated with the new energies. Thus, each plot in the following is a population average of the individual spectra for each of the two major conformers.

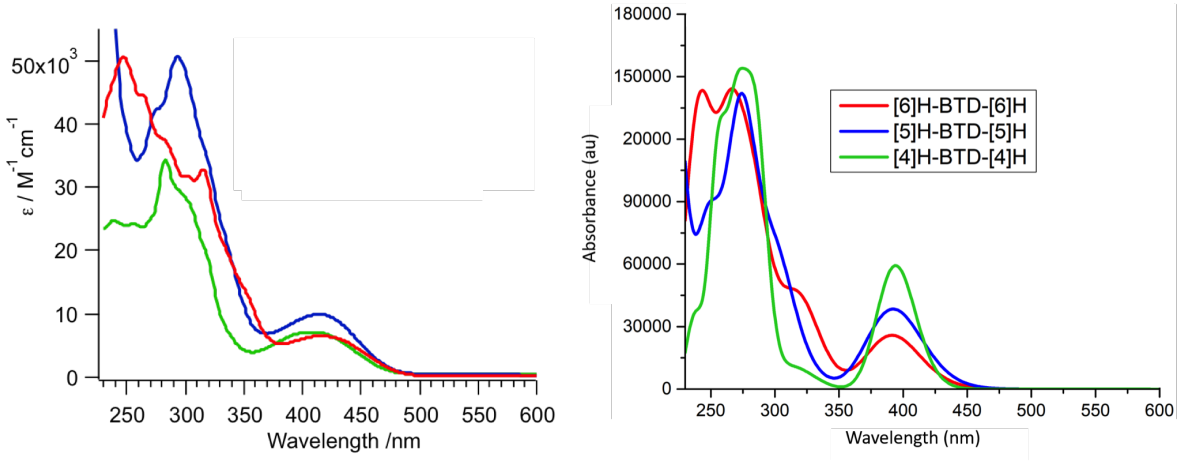


Figure 7.7: Experimental (left) and simulated (right) absorption spectra of M-[4]H-BTD-M-[4]H (green), M-[5]H-BTD-M-[5]H (blue), and M-[6]H-BTD-M-[6]H (red). Measurements were performed at room temperature in CH_2Cl_2 solutions ($2\text{-}3\cdot10^{-5}$ M) by the Avarvari group.

Absorption spectra are blue-shifted overall by about 20 nm but are otherwise in good agreement with experimental spectra (Figure 7.7). Compared to the other two molecules, the M-[4]H-BTD-M-[4]H peaks are slightly red shifted. M-[5]H-BTD-M-[5]H and M-[6]H-BTD-M-[6]H intensities are in good agreement with experimental spectra; however, the relative intensity compared to the other two molecules in the M-[4]H-BTD-M-[4]H spectrum is nearly two times larger than in experiment outside the 300-350 nm range. Fully matching all relative peak intensities would likely require more conformers for all three molecules, particularly M-[4]H-BTD-M-[4]H. Because there is an overall agreement in peak positions, we can use the simulated spectra for the peak assign-

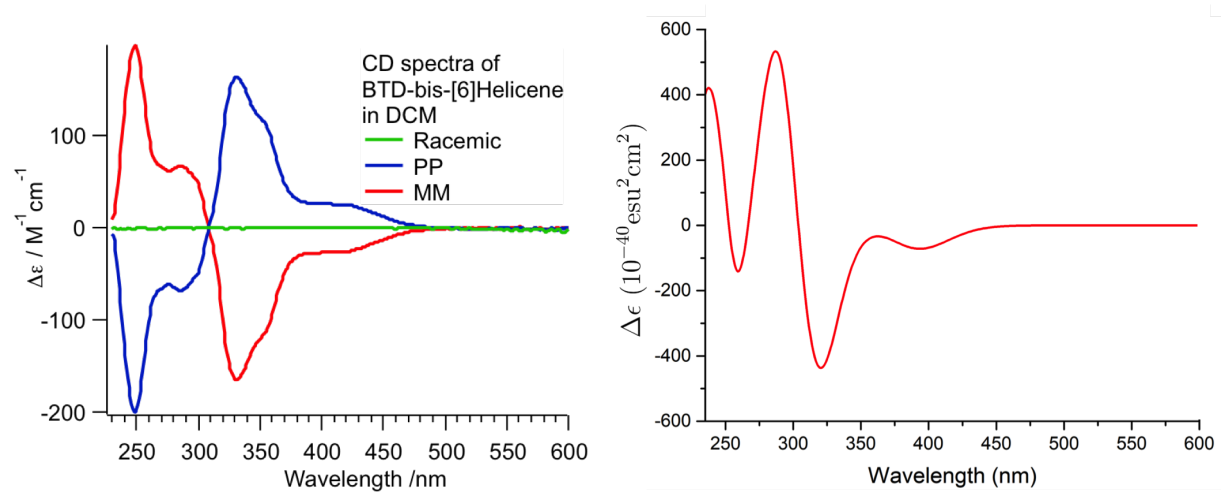


Figure 7.8: Experimental (left) and simulated (right) ECD spectra of M-[6]H-BTD-M-[6]H. Experiments are performed at room temperature in CH_2Cl_2 solutions (2-3*10^{-5}M) by the Avarvari group.

ment. Peak wavelengths in the following text are taken from the calculated results. Each peak is a combination of multiple states, $\pi \rightarrow \pi^*$ in nature, where the electron density moves across the helicene arms through the central BTD unit, and vice-versa. More specifically, the transitions can be categorized in five common types. In every molecule, the first peak near 400 nm is always a HOMO $\rightarrow$ LUMO charge transfer, where the HOMO is spread out across the arms of the helicene and the benzene unit of the BTD, and the LUMO is entirely on the BTD unit (**1**). The second common type of transition is a localized entirely on the helicene arms (**2**), and the third is localized entirely on the BTD unit (**3**). The fourth common transition (**4**) comprises states where the electron density is initially spread across both helicene arms and moves to the central BTD unit. The final common transition (**5**) is a charge transfer from the central BTD unit to the helicene arms.

In the M-[4]H-BTD-M-[4]H spectrum, there are peaks at 394, 320, 274, 259, and 237 nm. The 320 nm peak is a combination of type **2** and type **3** transitions, and the 274 nm peak is comprised of type **4** transitions. The 259 nm peak includes **2** transitions, and the 237 nm peak includes type **5** transitions. The peaks in the M-[5]H-BTD-M-[5]H spectrum are at 392, 304, 274, and 250 nm. The 304 nm peak is a type **2** transition, while the 274 nm peak is a combination of type **5** and type **3** transitions, and the 250 nm peak are type **4** transitions. Finally, there are peaks at 391, 317,

266, and 243 nm in the M-[6]H-BTD-M-[6]H spectrum. The 317 nm peak is characterized type **4** transitions, the 266 nm peak also includes type **4** transitions, and the 243 nm peak is a type **5** transition.

The ECD spectra for M-[6]H-BTD-M-[6]H are presented in Figure 7.8, and we find good agreement with the experimental spectrum. The main peaks are at 393, 330, 287, and 238 nm. The negative peak at 250 nm is likely an artifact due to the limitations in the conformational sampling of the calculations. As mentioned above, the 393 nm peak is a type **1** transition, the 330 nm peak is a combination of type **2** and type **4** transitions, the 287 nm peak includes type **4** and type **5** transitions. The final peak at 238 nm is a combination of type **3**, type **4**, and transitions where electron density begins on both helicene arms and moves to a single arm.

Figure 7.9 presents calculated and experimental emission spectra for all 3 molecules, which are in good agreement in both peak positions and relative intensities. The peak maximum of the M-[4]H-BTD-M-[4]H spectrum is correctly slightly blue shifted relative to the other two peaks, which are nearly on top of each other. All three transitions are type **1** $\pi^* \rightarrow \pi$ transitions.

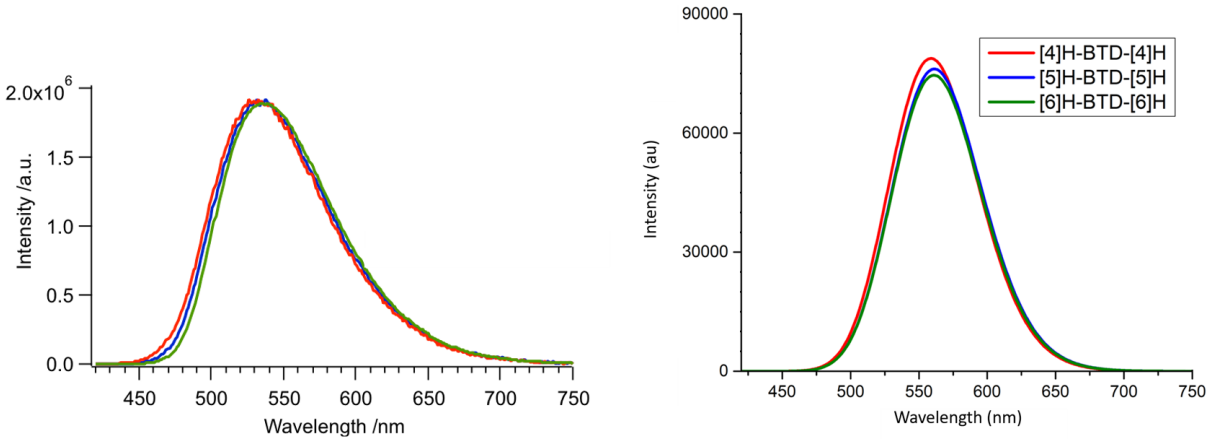


Figure 7.9: Experimental (left) and simulated (right) emission spectra of M-[4]H-BTD-M-[4]H (red), M-[5]H-BTD-M-[5]H (blue), and M-[6]H-BTD-M-[6]H (green). Measurements performed at room temperature in CH_2Cl_2 solutions (2-3*10^{-5}M) and λ_{ex} = 400 nm by the Avarvari group. Calculations are performed with CAM-B3LYP/aug-cc-pVDZ, in CH_2Cl_2 simulated using PCM and the SMD radii.

7.4 Discussion and Conclusions

Thiadiazole-fused [7]helicene **1** and [5]helicene **3**, together with their S-shaped double [4]helicene **2** and thiadiazole-benzanthracene **4** isomers, were investigated experimentally and with DFT simulations. Additionally, absorption, ECD and emission spectra series of M-[N]H-BTD-M[N]H helicenes are simulated and compared to experiments, and are found to be in good agreement. These results have allowed us to discuss the types of transitions which comprise the peaks in the spectra. The DFT calculations accurately reproduce the experimental spectra of **1**-**4**, and allow the assignment of the transitions between the thiadiazole and helicene units, which are $\pi \rightarrow \pi^*$ type. The calculations give a strong rationale for the fluorescence quenching of thiadiazole-[7]helicene **1**, which likely occurs by ISC from the minimum of the S1 state to one of the many triplet states. The shorter [5]helicene **3** is also not very emissive. In striking contrast, the S-shaped isomer **2** and benzanthracene **4** show strong luminescence.

The transitions involved in the absorption spectra of the M-[N]H-BTD-M[N]H molecules are all of $\pi \rightarrow \pi^*$ nature, and the accuracy of the calculations with respect to experimental results allows us to assign the peaks in each spectra. The transitions which make up the spectra can be subdivided into 5 basic types, which involve either some form of charge transfer between the central BTD unit or the helicene arms (transitions **1**, **4**, and **5**), or are localized on either the helicenes (transition **2**) or the BTD unit (transition **3**). These molecules are also highly emissive, in contrast to **1** and **3**.

These agreement between calculated and experimental results allows us to investigate enhancing the emission properties via functionalization of the helicene, and thus observation of circularly polarized luminescence, or analyze the electron-poor character in combination with electron-rich moieties to probe helical donor–acceptor systems. This would lead to better design of helicenes with targeted optical properties.

7.5 Acknowledgment

This work was supported in France by the CNRS through the GDR 3712 Chirafun, the University of Angers, and the French Ministry of Education and Research. The investigation was supported in part by the University of Kansas General Research Fund allocation #2302049, by the University of Kansas startup fund (T.A. and M.C.), the University of Geneva and by the Swiss National Science foundation (grant No 200020_152780). J.C. and N.A. warmly thank Chengshuo Shen (University of Rennes 1) and Flavia Pop (University of Angers) for technical help.

Chapter 8

Concluding Remarks

In this dissertation, we used electronic structure theory to examine the fundamental electronic effects that govern optical rotation, and developed tools to reduce the cost of $[\alpha]_\omega$ calculations. Our goal has been to help answer two fundamental questions that remain open in the study of optical rotation; namely, how can we relate the structure of a chiral molecule to the resulting magnitude and sign of $[\alpha]_\omega$ using chemical intuition beyond just performing black-box quantum mechanical calculations, and how does the environment affects this property?

We have worked to answer the first question by examining functional and structural effects that influence the intrinsic specific rotation of a prototypical class of chiral molecules: helicenes. We have also worked on the second question by studying whether implicit solvation models can reproduce experimental trends of solvation effects on $[\alpha]_\omega$ using a series of small organic molecules. Because accurate optical rotation calculations are computationally demanding,[2] we have also developed two approaches to reduce the computational cost: by selectively reducing the number of determinants used in the solution of the linear response equations, and by optimizing a small basis set specifically for this property.

In our study of the structure-property relationship of helicenes and $[\alpha]_\omega$, we have employed the $\tilde{S}_k$ method,[23,26,98] which decomposes $[\alpha]_\omega$ into occupied-virtual molecular orbital (MO) pair contributions, such that the orbitals which comprise the largest contributions can be visualized and analyzed. For instance, we have explored the effect of increasing the length and varying the pitch on unfunctionalized helicenes, and then proceeded to functionalize them with electron withdrawing and electron donating groups. We found that $[\alpha]_\omega$ can be described by two types of transitions, categorized by the orientation of the magnetic dipole with respect to the helical axis, and that the

addition of functional groups affects the specific rotation depending on their electron directing strength. Additionally, because canonical orbitals are delocalized across the entire molecule, and may be difficult to interpret, we have tested two localization schemes for the MO analysis of $[\alpha]_\omega$, i.e., the Boys,[111,112] and of Pipek & Mezey methods.[113] We found that localized MOs can indeed be useful when the chromophoric groups in the molecules are small.

In order to study solvent effects on $[\alpha]_\omega$, we collaborated with the Vaccaro group at Yale University, who performed gas and solution phase optical rotation measurements.[45,219] Using rigid chiral molecules to limit conformational effects, we compared theoretical and experimental values of $[\alpha]_\omega$ in gas phase and in a series of solvents of increasing polarity ranging from cyclohexane to acetonitrile. We found that, while gas phase calculations are capable of reproducing experimental values, continuum solvation models fail to reproduce both solvent shifts and relative trends in the solvent effect, even when including vibrational corrections.

The effort to reduce the computational cost of the linear response equations necessary to compute $[\alpha]_\omega$ was based on our previous work on the MO decomposition of the optical rotation tensor, where we showed that many MO pairs contributed very little to the final value of the property.[23,26,98] Because solving the linear response equations is an iterative procedure, we have used the guess density to find and discard MOs with small contributions. We have tested a number of selection criteria and cutoff values, using 51 molecules, two basis sets, and two functionals for the assessment. This method accurately reproduced $[\alpha]_\omega$ calculations despite discarding a significant number of excited determinants, which resulted in significant computational savings. The results mainly depended on the magnitude of the cutoff, and only moderately on the selection criteria.

We have also built a small specialized basis set inspired by the work of Wiberg *et al.*, who showed that diffuse functions have the largest effect on accurate $[\alpha]_\omega$ calculations.[61] The basis sets were built by combining the 3-21G basis set with the diffuse functions of either the aug-cc-pVDZ and aug-cc-pVTZ basis sets: augD-3-21G and augT3-3-21G, respectively. We then optimized the exponents of the diffuse functions to provide the best agreement (in a least mean square sense) with $[\alpha]_D$ computed with CAM-B3LYP and the corresponding full basis set for a training set of

21 organic molecules. Each basis set was then tested at different frequencies, with a different functional, and with a different set of molecules. The new basis sets proved to be successful at reproducing the $[\alpha]_\omega$ values of the target basis sets. The errors were on average small across the training and control sets at all tested wavelengths and functionals, and augT3-3-21G was as effective as aug-cc-pVDZ in reproducing aug-cc-pVTZ calculations. Timing data showed that the new basis sets were 3-4 times as fast as aug-cc-pVDZ calculations, and more than 20 times faster than aug-cc-pVTZ calculations.

Finally, we have also worked with the Avarvari group at the Université d'Angers on a joint experimental-computational study of functionalized helicenes, because of their interesting electronic properties. We simulated absorption, emission, and ECD spectra, and characterized the transitions. For instance, the simulations allowed us to determine that the cause of fluorescence quenching, observed experimentally, is the abundance of triplet states available for intersystem crossing from singlet excited states.[81]

Our efforts over the last 5 years in the study of chiral molecules has increased our understanding of environmental effects, and the structure-property relationship between chiral molecules and OR. However, further work is necessary. For instance, the failures of continuum solvation models suggests that solvation effects are likely not just electrostatic in nature, and should be investigated with fully quantum mechanical calculations of a chiral molecule with the surrounding first and maybe second solvation shell, perhaps in concert with the smaller basis sets detailed here. Additionally, the orbital localization analysis should be employed on more molecules with compact chromophores as localized orbitals can provide an easier to interpret method to study the effect of functionalization and structural changes. This analysis, may be also used to analyze quantum mechanical solute-solvent interactions from the QM solvation shell calculations, as the effect of specific solute-solvent interactions on OR are still not well understood. The orbital windowing technique should be implemented in a efficient code to actually reduce the cost of production level calculations, and it could be used in concert with the new reduced basis sets we developed. The same approach should then be extended to CC methods, so that accurate calculations can be per-

formed on larger molecules.[14]

While the coming years may not see the ability to predict the $[\alpha]_\omega$ of a given structure without quantum calculations, our work will assist in the design of chiral molecules. With efficient implementations of our cost saving methods and usage of our technique to decompose the OR tensor, functionalization and conformation studies may give experimental groups more control over the development of chiral molecules with desired optical properties. We believe we may soon see the answer to the question of how solute-solvent specific interactions produce such large changes in the OR. Cheaper calculations will allow us to explore larger solute-solvent systems, while our decomposition tool will allow us to probe these interactions.

Chapter 9

References

[1] National Center for Biotechnology Information. PubChem Database. D-Limonene. `https://pubchem.ncbi.nlm.nih.gov/compound/D-Limonene`, Accessed: 07-16-2019.

[2] Vaccaro, P. H. *Comprehensive Chiroptical Spectroscopy, Instrumentation, Methodologies, and Theoretical Simulations Vol 1*; John Wiley & Sons, 2012; pp 265–324, Berova; Polavarapu; Nakanishi; Woody.

[3] Vargesson, N. Thalidomide-induced teratogenesis: History and mechanisms. *Birth Defects Res. C: Embryo Today* **2015**, *105*, 140–156.

[4] Harada, N.; Watanabe, M.; Kuwahara, S. In *Chiral Analysis*; Busch, K. W., Busch, M. A., Eds.; Elsevier: Amsterdam, 2006; pp 661 – 692.

[5] Polavarapu, P. L. Ab initio molecular optical rotations and absolute configurations. *Mol. Phys.* **1997**, *91*, 551–554.

[6] Cheeseman, J. R.; Frisch, M. J.; Devlin, F. J.; Stephens, P. J. Hartree-Fock and density functional theory ab initio calculation of optical rotation using GIAOs: basis set dependence. *J. Phys. Chem. A* **2000**, *104*, 1039–1046.

[7] Grimme, S. Calculation of frequency dependent optical rotation using density functional response theory. *Chem. Phys. Lett.* **2001**, *339*, 380–388.

[8] Polavarapu, P. L. Optical rotation: recent advances in determining the absolute configuration. *Chirality* **2002**, *14*, 768–781.

[9] Autschbach, J.; Patchkovskii, S.; Ziegler, T.; van Gisbergen, S. J. a.; Jan Baerends, E. Chiroptical properties from time-dependent density functional theory. II. Optical rotations of small to medium sized organic molecules. *J. Chem. Phys.* **2002**, *117*, 581–592.

[10] Ruud, K.; Helgaker, T. Optical rotation studied by density-functional and coupled-cluster methods. *Chem. Phys. Lett.* **2002**, *352*, 533–539.

[11] Stephens, P. J.; McCann, D. M.; Cheeseman, J. R.; Frisch, M. J. Determination of absolute configurations of chiral molecules using ab initio time-dependent density functional theory calculations of optical rotation: How reliable are absolute configurations obtained for molecules with small rotations? *Chirality* **2005**, *17*, 52–64.

[12] Grimme, S.; Bahlmann, A.; Haufe, G. Ab initio calculations for the optical rotations of conformationally flexible molecules: A case study on six-, seven-, and eight-membered fluorinated cycloalkanol esters. *Chirality* **2002**, *14*, 793–797.

[13] Ruud, K.; Stephens, P. J.; Devlin, F. J.; Taylor, P. R.; Cheeseman, J. R.; Frisch, M. J. Coupled-cluster calculations of optical rotation. *Chem. Phys. Lett.* **2003**, *373*, 606–614.

[14] Tam, M. C.; Russ, N. J.; Crawford, T. D. Coupled cluster calculations of optical rotatory dispersion of (S)-methyloxirane. *J. Chem. Phys.* **2004**, *121*, 3550–3557.

[15] Crawford, T. D. Ab initio calculation of molecular chiroptical properties. *Theor. Chem. Acc.* **2006**, *115*, 227–245.

[16] Crawford, T. D.; Stephens, P. J. Comparison of time-dependent density-functional theory and coupled cluster theory for the calculation of the optical rotations of chiral molecules. *J. Phys. Chem. A.* **2008**, *112*, 1339–1345.

[17] Pedersen, T. B.; Koch, H.; Boman, L.; Sánchez De Merás, A. M. J. Origin invariant calculation of optical rotation without recourse to London orbitals. *Chem. Phys. Lett.* **2004**, *393*, 319–326.

[18] Crawford, T. D.; Owens, L. S.; Tam, M. C.; Schreiner, P. R.; Koch, H. Ab initio calculation of optical rotation in (P)-(+)-[4]triangulane. *J. Am. Chem. Soc.* **2005**, *127*, 1368–1369.

[19] Krykunov, M.; Autschbach, J. Calculation of optical rotation with time-periodic magnetic-field-dependent basis functions in approximate time-dependent density-functional theory. *J. Chem. Phys.* **2005**, *123*, 114103.

[20] Autschbach, J. Computing chiroptical properties with first-principles theoretical methods: background and illustrative examples. *Chirality* **2009**, *21*, E116–E152.

[21] Moore, B.; Srebro, M.; Autschbach, J. Analysis of optical activity in terms of bonds and lone-pairs: The exceptionally large optical rotation of norbornenone. *J. Chem. Theory Comput.* **2012**, *8*, 4336–4346.

[22] Caricato, M.; Vaccaro, P. H.; Crawford, T. D.; Wiberg, K. B.; Lahiri, P. Insights on the origin of the unusually large specific rotation of (1S,4S)-Norbornenone. *J. Phys. Chem. A* **2014**, *118*, 4863–4871.

[23] Caricato, M. Orbital analysis of molecular optical activity based on configuration rotatory strength. *J. Chem. Theory Comput.* **2015**, *11*, 1349–1353.

[24] Murphy, V. L.; Kahr, B. Hückel theory and optical activity. *J. Am. Chem. Soc.* **2015**, *137*, 5177–5183.

[25] Kondru, R. K.; Wipf, P.; Beratan, D. N. Atomic Contributions to the Optical Rotation Angle as a Quantitative Probe of Molecular Chirality. *Science* **1998**, *282*, 2247–2250.

[26] Caricato, M. Conformational effects on specific rotation: A theoretical study based on the $\tilde{S}_k$ method. *J. Phys. Chem. A* **2015**, *119*, 8303–8310.

[27] Lahiri, P.; Wiberg, K. B.; Vaccaro, P. H. Intrinsic Optical Activity and Large-Amplitude Displacement: Conformational Flexibility in (R)-Glycidyl Methyl Ether. *J. Phys. Chem. A* **2015**, *119*, 8311–8327.

[28] Tam, M. C.; Crawford, T. D.; Tech, V. Ab Initio Determination of Optical Rotatory Dispersion in the Conformationally Flexible. *J. Phys. Chem. A* **2006**, *110*, 2290–2298.

[29] Wiberg, K. B.; Wang, Y.; Vaccaro, P. H.; Cheeseman, J. R.; Luderer, M. R. Conformational effects on optical rotation: 2-substituted butanes. *J. Am. Chem. Soc.* **2005**, *109*, 3405–3410.

[30] Wiberg, K. B.; Vaccaro, P. H.; Cheeseman, J. R. Conformational Effects on Optical Rotation. 3-Substituted 1-Butenes. *J. Am. Chem. Soc.* **2003**, *125*, 1888–1896.

[31] Lahiri, P.; Wiberg, K. B.; Vaccaro, P. H. A tale of two carenes: Intrinsic optical activity and large-amplitude nuclear displacement. *J. Phys. Chem. A* **2012**, *116*, 9516–9533.

[32] Wiberg, K. B.; Wang, Y. G.; Vaccaro, P. H.; Cheeseman, J. R.; Trucks, G.; Frisch, M. J. Optical activity of 1-butene, butane, and related hydrocarbons. *J. Phys. Chem. A* **2004**, *108*, 32–38.

[33] Pedersen, T. B.; Kongsted, J.; Crawford, T. D.; Ruud, K. On the importance of vibrational contributions to small-angle optical rotation: Fluoro-oxirane in gas phase and solution. *J. Chem. Phys.* **2009**, *130*.

[34] Lahiri, P.; Wiberg, K. B.; Vaccaro, P. H. Intrinsic optical activity and conformational flexibility: The role of size-dependent ring morphology in model cycloketones. *J. Phys. Chem. A* **2013**, *117*, 12382–12400.

[35] Crawford, T. D.; Tam, M. C.; Abrams, M. L. The problematic case of (S)-methylthiirane: electronic circular dichroism spectra and optical rotatory dispersion. *Molecular Physics* **2007**, *105*, 2607–2617.

[36] Mort, B. C.; Autschbach, J. Temperature Dependence of the Optical Rotation of Fenchone Calculated by Vibrational Averaging. *J. Phys. Chem. A* **2006**, *110*, 11381–11383.

[37] Mort, B. C.; Autschbach, J. Temperature dependence of the optical rotation in six bicyclic organic molecules calculated by vibrational averaging. *ChemPhysChem* **2007**, *8*, 605–616.

[38] Kongsted, J.; Pedersen, T. B.; Jensen, L.; Hansen, A. E.; Mikkelsen, K. V. Coupled cluster and density functional theory study of the vibrational contribution to the optical rotation of (\textit{S})-propylene oxide. *J. Am. Chem. Soc.* **2006**, *128*, 976–982.

[39] Ruud, K.; Åstrand, P. O.; Taylor, P. R. An efficient approach for calculating vibrational wave functions and zero-point vibrational corrections to molecular properties of polyatomic molecules. *J. Chem. Phys.* **2000**, *112*, 2668–2683.

[40] Åstrand, P.-O.; Ruud, K.; Taylor, P. R. Calculation of the vibrational wave function of polyatomic molecules. *J. Chem. Phys.* **2000**, *112*, 2655.

[41] Ruud, K.; Taylor, P. R.; Åstrand, P. O. Zero-point vibrational effects on optical rotation. *Chem. Phys. Lett.* **2001**, *337*, 217–223.

[42] Mort, B. C.; Autschbach, J. Magnitude of zero-point vibrational corrections to optical rotation in rigid organic molecules: A time-dependent density functional study. *J. Phys. Chem. A* **2005**, *109*, 8617–8623.

[43] Pedersen, T. B.; Kongsted, J.; Crawford, T. D. Gas Phase Optical Rotation Calculated from Coupled Cluster Theory with Zero-Point Vibrational Corrections from Density Funcional Theory. *Chirality* **2009**, *21*, E68–E75.

[44] Neugebauer, J.; Jan Baerends, E.; Nooijen, M.; Autschbach, J. Importance of vibronic effects on the circular dichroism spectrum of dimethyloxirane. *J. Chem. Phys.* **2005**, *122*.

[45] Müller, T.; Wiberg, K. B.; Vaccaro, P. H. Cavity ring-down polarimetry (CRDP): A new scheme for probing circular birefringence and circular dichroism in the gas phase. *J. Phys. Chem. A* **2000**, *104*, 5959–5968.

[46] Mennucci, B.; Tomasi, J.; Cammi, R.; Cheeseman, J. R.; Frisch, M. J.; Devlin, F. J.; Gabriel, S.; Stephens, P. J. Polarizable continuum model (PCM) calculations of solvent effects on optical rotations of chiral molecules. *J. Phys. Chem. A* **2002**, *106*, 6102–6113.

[47] da Silva, C. O.; Mennucci, B.; Vreven, T. Density functional study of the optical rotation of glucose in aqueous solution. *J. Org. Chem.* **2004**, *69*, 8161–4.

[48] Stephens, P. J.; Devlin, F. J.; Cheeseman, J. R.; Frisch, M. J.; Mennucci, B.; Tomasi, J. Prediction of optical rotation using density functional theory: 6,8-dioxabicyclo[3.2.1]octanes. *The J. Phys. Chem. A* **2000**, *11*, 2443.

[49] Caricato, M. Implementation of the CCSD-PCM linear response function for frequency dependent properties in solution: Application to polarizability and specific rotation. *J. Chem. Phys.* **2013**, *139*, 114103.

[50] Fischer, A. T.; Compton, R. N.; Pagni, R. M. Solvent effects on the optical rotation of (S)-(-)-alpha-methylbenzylamine. *J. Phys. Chem. A* **2006**, *110*, 7067–7071.

[51] Wilson, S. M.; Wiberg, K. B.; Murphy, M. J.; Vaccaro, P. H. The effects of conformation and solvation on optical rotation: Substituted epoxides. *Chirality* **2008**, *20*, 357–369.

[52] Lahiri, P.; Wiberg, K. B.; Vaccaro, P. H. Dispersive Optical Activity of (R)-Methylene Norbornene: Intrinsic Response and Solvation Effects. *J. Phys. Chem. A* **2017**,

[53] Marchesan, D.; Coriani, S.; Forzato, C.; Nitti, P.; Pitacco, G.; Ruud, K. Optical rotation calculation of a highly flexible molecule: The case of paraconic acid. *J. Phys. Chem. A* **2005**, *109*, 1449–1453.

[54] Crawford, T. D.; Kumar, A.; Hannon, K. P.; Höfener, S.; Visscher, L. Frozen-Density Embedding Potentials and Chiroptical Properties. *J. Chem. Theory Comput.* **2015**, *11*, 5305–5315.

[55] Kundrat, M. D.; Autschbach, J. Ab initio and density functional theory modeling of the chiroptical response of glycine and alanine in solution using explicit solvation and molecular dynamics. *J. Chem. Theory Comput.* **2008**, *4*, 1902–1914.

[56] Kundrat, M. D.; Autschbach, J. Modeling of the chiroptical response of chiral amino acids in solution using explicit solvation and molecular dynamics. *J. Chem. Theory Comput.* **2009**, *5*, 1051–1060.

[57] Mukhopadhyay, P.; Zuber, G.; Wipf, P.; Beratan, D. N. Contribution of a solute's chiral solvent imprint to optical rotation. *Angew. Chem. Int. Ed.* **2007**, *46*, 6450–6452.

[58] Mukhopadhyay, P.; Zuber, G.; Goldsmith, M.-R.; Wipf, P.; Beratan, D. N. Solvent Effect on Optical Rotation: a Case Study of Methyloxirane in Water. *ChemPhysChem* **2006**, *7*, 2483–2486.

[59] Cappelli, C. Integrated QM/polarizable MM/continuum approaches to model chiroptical properties of strongly interacting solute–solvent systems. *Int. J. Quantum Chem.* **2016**, *116*, 1532–1542.

[60] Egidi, F.; Giovannini, T.; Del Frate, G.; Lemler, P. M.; Vaccaro, P. H.; Cappelli, C. A combined experimental and theoretical study of optical rotatory dispersion for (R)-glycidyl methyl ether in aqueous solution. *Phys. Chem. Chem. Phys* **2019**, *21*, 3644–3655.

[61] Wiberg, K. B.; Caricato, M.; Wang, Y.-g.; Vaccaro, P. H. Towards the accurate and efficient calculation of optical rotatory dispersion using augmented minimal basis sets. *Chirality* **2013**, *25*, 606–616.

[62] Baranowska-Łączkowska, A.; Łączkowski, K. Z. The ORP basis set designed for optical rotation calculations. *J. Comput. Chem.* **2013**, *34*, 2006–2013.

[63] Baranowska-Łączkowska, A.; Łączkowski, K. Z.; Henriksen, C.; Fernández, B.; Kozak, M.; Zielińska, S. New basis set for the prediction of the specific rotation in flexible biological molecules. *RSC Adv.* **2016**, *6*, 19897–19902.

[64] Russ, N. J.; Crawford, T. D. Local correlation domains for coupled cluster theory: optical rotation and magnetic-field perturbations. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3345–3352.

[65] McAlexander, H. R.; MacH, T. J.; Crawford, T. D. Localized optimized orbitals, coupled cluster theory, and chiroptical response properties. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7830–7836.

[66] McAlexander, H. R.; Crawford, T. D. A Comparison of Three Approaches to the Reduced-Scaling Coupled Cluster Treatment of Non-Resonant Molecular Response Properties. *J. Chem. Theory Comput.* **2016**, *12*, 209–222.

[67] Gingras, M. One hundred years of helicene chemistry. Part 1: Non-stereoselective syntheses of carbohelicenes. *Chem. Soc. Rev.* **2013**, *42*, 968–1006.

[68] Gingras, M.; Félix, G.; Peresutti, R. One hundred years of helicene chemistry. Part 2: Stereoselective syntheses and chiral separations of carbohelicenes. *Chem. Soc. Rev.* **2013**, *42*, 1007–1050.

[69] Gingras, M. One hundred years of helicene chemistry. Part 3: Applications and properties of carbohelicenes. *Chem. Soc. Rev.* **2013**, *42*, 1051–1095.

[70] Yang, Y.; Da Costa, R. C.; Smilgies, D. M.; Campbell, A. J.; Fuchter, M. J. Induction of circularly polarized electroluminescence from an achiral light-emitting polymer via a chiral small-molecule dopant. *Adv. Mater.* **2013**, *25*, 2624–2628.

[71] Brandt, J. R.; Wang, X.; Yang, Y.; Campbell, A. J.; Fuchter, M. J. Circularly Polarized Phosphorescent Electroluminescence with a High Dissymmetry Factor from PHOLEDs Based on a Platinahelicene. *J. Am. Chem. Soc.* **2016**, *138*, 9743–9746.

[72] Biet, T.; Cauchy, T.; Sun, Q.; Ding, J.; Hauser, A.; Oulevey, P.; Bürgi, T.; Jacquemin, D.; Vanthuyne, N.; Crassous, J.; Avarvari, N. Triplet state CPL active helicene-dithiolene platinum bipyridine complexes. *ChemComm* **2017**, *53*, 9210–9213.

[73] Yang, Y.; Da Costa, R. C.; Fuchter, M. J.; Campbell, A. J. Circularly polarized light detection by a chiral organic semiconductor transistor. *Nat. Photonics* **2013**, *7*, 634–638.

[74] Josse, P.; Favereau, L.; Shen, C.; Dabos-Seignon, S.; Blanchard, P.; Cabanetos, C.; Crassous, J. Enantiopure versus Racemic Naphthalimide End-Capped Helicenic Non-fullerene Electron Acceptors: Impact on Organic Photovoltaics Performance. *Chem. Eur. J.* **2017**, *23*, 6277–6281.

[75] Hatakeyama, T.; Hashimoto, S.; Oba, T.; Nakamura, M. Azaboradibenzo[6]helicene: Carrier inversion induced by helical homochirality. *J. Am. Chem. Soc.* **2012**, *134*, 19600–19603.

[76] Salerno, F.; Rice, B.; Schmidt, J. A.; Fuchter, M. J.; Nelson, J.; Jelfs, K. E. The influence of nitrogen position on charge carrier mobility in enantiopure aza[6]helicene crystals. *Phys. Chem. Chem. Phys.* **2019**, *21*, 5059–5067.

[77] Yang, Y.; Rice, B.; Shi, X.; Brandt, J. R.; Correa Da Costa, R.; Hedley, G. J.; Smilgies, D. M.; Frost, J. M.; Samuel, I. D.; Otero-De-La-Roza, A.; Johnson, E. R.; Jelfs, K. E.; Nelson, J.; Campbell, A. J.; Fuchter, M. J. Emergent Properties of an Organic Semiconductor Driven by its Molecular Chirality. *ACS Nano* **2017**, *11*, 8329–8338.

[78] Isla, H.; Crassous, J. Helicene-based chiroptical switches. *C. R. Chim.* **2016**, *19*, 39–49.

[79] Saleh, N.; Vanthuyne, N.; Bonvoisin, J.; Autschbach, J.; Srebro-Hooper, M.; Crassous, J. Redox-triggered chiroptical switching activity of ruthenium(III)-bis-(β-diketonato) complexes bearing a bipyridine-helicene ligand. *Chirality* **2018**, *30*, 592–601.

[80] Srebro, M.; Anger, E.; Moore, B.; Vanthuyne, N.; Roussel, C.; Réau, R.; Autschbach, J.; Crassous, J. Ruthenium-Grafted Vinylhelicenes: Chiroptical Properties and Redox Switching. *Chem. Eur. J.* **2015**, *21*, 17100–17115.

[81] Biet, T.; Martin, K.; Hankache, J.; Hellou, N.; Hauser, A.; Bürgi, T.; Vanthuyne, N.; Aharon, T.; Caricato, M.; Crassous, J.; Avarvari, N. Triggering emission with the helical turn in thiadiazole-helicenes. *Chem. Eur. J.* **2017**, *23*, 437–446.

[82] Shen, C.; Anger, E.; Srebro, M.; Vanthuyne, N.; Toupet, L.; Roussel, C.; Autschbach, J.; Réau, R.; Crassous, J. Diastereo- and enantioselective synthesis of organometallic bis(helicene)s by a combination of C-H activation and dynamic isomerization. *Chem. Eur. J.* **2013**, *19*, 16722–16728.

[83] Bouvier, R.; Durand, R.; Favereau, L.; Srebro-Hooper, M.; Dorcet, V.; Roisnel, T.; Vanthuyne, N.; Vesga, Y.; Donnelly, J.; Hernandez, F.; Autschbach, J.; Trolez, Y.; Crassous, J. Helicenes grafted with 1,1,4,4-tetracyanobutadiene moieties: π-helical push-pull systems with strong electronic circular dichroism and two-photon absorption. *Chem. Eur. J.* **2018**, *24*, 14484–14494.

[84] Hellou, N.; Jahier-Diallo, C.; Baslé, O.; Srebro-Hooper, M.; Toupet, L.; Roisnel, T.; Caytan, E.; Roussel, C.; Vanthuyne, N.; Autschbach, J.; Mauduit, M.; Crassous, J. Electronic and chiroptical properties of chiral cycloiridiated complexes bearing helicenic NHC ligands. *ChemComm* **2016**, *52*, 9243–9246.

[85] Shen, C.; Loas, G.; Srebro-Hooper, M.; Vanthuyne, N.; Toupet, L.; Cador, O.; Paul, F.; López Navarrete, J. T.; Ramírez, F. J.; Nieto-Ortega, B.; Casado, J.; Autschbach, J.; Vallet, M.; Crassous, J. Iron alkynyl helicenes: Redox-triggered chiroptical tuning in the IR and near-IR spectral regions and suitable for telecommunications applications. *Angew. Chem. Int. Ed.* **2016**, *55*, 8062–8066.

[86] Isla, H.; Srebro-Hooper, M.; Jean, M.; Vanthuyne, N.; Roisnel, T.; Lunkley, J. L.; Muller, G.; Williams, J. A.; Autschbach, J.; Crassous, J. Conformational changes and chiroptical switching of enantiopure bis-helicenic terpyridine upon Zn2+ binding. *ChemComm* **2016**, *52*, 5932–5935.

[87] Dhbaibi, K.; Favereau, L.; Srebro-Hooper, M.; Jean, M.; Vanthuyne, N.; Zinna, F.; Jamoussi, B.; Di Bari, L.; Autschbach, J.; Crassous, J. Exciton coupling in

diketopyrrolopyrrole-helicene derivatives leads to red and near-infrared circularly polarized luminescence. *Chem. Sci.* **2018**, *9*, 735–742.

[88] Bensalah-Ledoux, A.; Pitrat, D.; Reynaldo, T.; Srebro-Hooper, M.; Moore, B.; Autschbach, J.; Crassous, J.; Guy, S.; Guy, L. Large-scale synthesis of helicene-like molecules for the design of enantiopure thin films with strong chiroptical activity. *Chem. Eur. J.* **2016**, *22*, 3333–3346.

[89] Mukhopadhyay, A.; Hossen, T.; Ghosh, I.; Koner, A. L.; Nau, W. M.; Sahu, K.; Moorthy, J. N. Helicity-Dependent Regiodifferentiation in the Excited-State Quenching and Chiroptical Properties of Inward/Outward Helical Coumarins. *Chem. Eur. J.* **2017**, *23*, 14797–14805.

[90] Barron, L. D. *Molecular Light Scattering and Optical Activity*, 2nd ed.; Cambridge University Press, 2004.

[91] Rosenfeld, L. Quantenmechanische Theorie der natürlichen optischen Aktivität von Flüssigkeiten und Gasen. *Z. Phys.* **1929**, *52*, 161–174.

[92] Wiberg, K. B.; Wang, Y. G.; Wilson, S. M.; Vaccaro, P. H.; Cheeseman, J. R. Sum-over-states calculation of the specific rotations of some substituted oxiranes, chloropropionitrile, ethane, and norbornenone. *J. Phys. Chem. A* **2006**, *110*, 13995–14002.

[93] Crawford, T. D.; Tam, M. C.; Abrams, M. L. The current state of ab initio calculations of optical rotation and electronic circular dichroism spectra. *J. Phys. Chem. A* **2007**, *111*, 12057–68.

[94] London, F. Théorie quantique des courants interatomiques dans les combinaisons aromatiques. *J. Phys. Radium* **1937**, *8*, 397–409.

[95] Ditchfield, R. Self-consistent perturbation theory of diamagnetism. *Mol. Phys.* **1974**, *27*, 789–807.

[96] Helgaker, T.; Ruud, K.; Bak, K. L.; Jørgensen, P.; Olsen, J. Vibrational Raman optical activity calculations using London atomic orbitals. *Faraday Discuss.* **1994**, *99*, 165–180.

[97] Bak, K. L.; Hansen, A. E.; Ruud, K.; Helgaker, T.; Olsen, J.; Jørgensen, P. Ab initio calculation of electronic circular dichroism fortrans-cyclooctene using London atomic orbitals. *Theor. Chem. Acc.* **1995**, *90*, 441–458.

[98] Aharon, T.; Caricato, M. Configuration Space Analysis of the Specific Rotation of Helicenes. *J. Phys. Chem. A* **2019**, *123*, 4406–4418.

[99] Murphy, V. L.; Farfan, C.; Kahr, B. Chiroptical structure - property relations in cyclo[18]carbon and its in silico hydrogenation products. *Chirality* **2018**, *30*, 325–331.

[100] Wiberg, K. B. Chirality induced by the interaction of C=C and C=X bonds (X=CH2, NH, NH2+, O, and S) separated by a methylene group. *J. Phys. Chem. A* **2016**, *120*, 7771–7777.

[101] Wiberg, K. B. Chiroptical properties of imines derived from *R*-(+)-norbornenone: The role of electronegativity differences. *The J. Phys. Chem. A* **2017**, *121*, 8247–8250.

[102] Furche, F.; Ahlrichs, R.; Wachsmann, C.; Weber, E.; Sobanski, A.; Vögtle, F.; Grimme, S. Circular dichroism of helicenes investigated by time-dependent density functional theory. *J. Am. Chem. Soc.* **2000**, *122*, 1717–1724.

[103] Verbiest, T.; Verbiest, T.; Elshocht, S. V.; Kauranen, M.; Hellemans, L.; Snauwaert, J.; Nuckolls, C.; Katz, T. J. Strong enhancement of nonlinear optical properties through supramolecular chirality. *Science* **1998**, *282*, 913–915.

[104] Bossi, A.; Licandro, E.; Maiorana, S.; Rigamonti, C.; Righetto, S.; Stephenson, G. R.; Spassova, M.; Botek, E.; Champagne, B. Theoretical and experimental investigation of electric field induced second harmonic generation in tetrathia[7]helicenes. *J. Phys. Chem. C* **2008**, *112*, 7900–7907.

[105] Shen, C.; He, X.; Toupet, L.; Norel, L.; Rigaut, S.; Crassous, J. Dual redox and optical control of chiroptical activity in photochromic dithienylethenes decorated with hexahelicene and bis-ethynyl-ruthenium units. *Organometallics* **2018**, *37*, 697–705.

[106] Hellou, N.; Macé, A.; Martin, C.; Dorcet, V.; Roisnel, T.; Jean, M.; Vanthuyne, N.; Berrée, F.; Carboni, B.; Crassous, J. Synthesis of carbo[6]helicene derivatives grafted with amino or aminoester substituents from enantiopure [6]helicenyl boronates. *J. Org. Chem.* **2018**, *83*, 484–490.

[107] OuYang, J.; Crassous, J. Chiral multifunctional molecules based on organometallic helicenes: Recent advances. *Coord. Chem. Rev.* **2018**, *376*, 533–547.

[108] Nakai, Y.; Mori, T.; Inoue, Y. Circular dichroism of (di)methyl- and diaza[6]helicenes. A combined theoretical and experimental study. *J. Phys. Chem. A* **2013**, *117*, 83–93.

[109] Nakai, Y.; Mori, T.; Inoue, Y. Theoretical and experimental studies on circular dichroism of carbo[N]helicenes. *J. Phys. Chem. A* **2012**, *116*, 7372–7385.

[110] Tanaka, H.; Kato, Y.; Fujiki, M.; Inoue, Y.; Mori, T. Combined experimental and theoretical study on circular dichroism and circularly polarized luminescence of configurationally robust D3-symmetric triple pentahelicene. *J. Phys. Chem. A* **2018**, 7378–7384.

[111] Foster, J.; Boys, S. F. Canonical configuration interaction procedure. *Rev. Mod. Phys.* **1960**, *32*, 300–302.

[112] Boys, S. F. Construction of some molecular orbitals to be approximately invariant for changes from one molecule to another. *Rev. Mod. Phys.* **1960**, *32*, 296–299.

[113] Pipek, J.; Mezey, P. G. A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. *J. Chem. Phys.* **1989**, *90*, 4916–4926.

[114] Frisch, M. J. et al. Gaussian Development Version. 2009.

[115] Yanai, T.; Tew, D. P.; Handy, N. C. A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57.

[116] Woon, D. E.; Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. IV. Calculation of static electrical response properties. *J. Chem. Phys.* **1994**, *100*, 2975–2988.

[117] Becke, A. D. A new mixing of Hartree-Fock and local density functional theories. *J. Chem. Phys.* **1993**, *98*, 1372–1377.

[118] Becke, A. D. Density functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

[119] Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789.

[120] Martin, R. H.; Flammang-Barbieux, M.; Cosyn, J. P.; Gelbcke, M. 1-Synthesis of octa- and nonahelicenes. 2-New syntheses of hexa- and heptahelicenes. 3-Optical rotation and O.R.D. of heptahelicene. *Tetrahedron Lett.* **1968**, *9*, 3507–3510.

[121] Martin, R. H.; Libert, V. J. Helicenes - the use of resolved hexahelicene-2-carboxylic acid as a common precurser for the photochemical synthesis of optically pure octa-helicenes, nona-helicenes, deca-helicenes, undeca-helicenes and trideca-helicenes - thermal racemization of deca-helicenes and undeca helicenes. *J. Chem. Res. (M)* **1980**, *4*, 1940–1950.

[122] Besler, B. H.; Merz, K. M.; Kollman, P. A. Atomic charges derived from semiempirical methods. *J. Comput. Chem.* **1990**, *11*, 431–439.

[123] Singh, U. C.; Kollman, P. A. An approach to computing electrostatic charges for molecules. *J. Comput. Chem.* **1984**, *5*, 129–145.

[124] Lowry, T. *Optical Rotatory Power*; Longmans, Green and Co.: London, 1935; Donnan, FG.

[125] Cheeseman, J. R.; Frisch, M. J.; Devlin, F. J.; Stephens, P. J. Hartree−Fock and Density Functional Theory ab Initio Calculation of Optical Rotation Using GIAOs: basis Set Dependence. *J. Phys. Chem. A.* **2000**, *104*, 1039–1046.

[126] Wilson, S. M.; Wiberg, K. B.; Cheeseman, J. R.; Frisch, M. J.; Vaccaro, P. H. Nonresonant Optical Activity of Isolated Organic Molecules. *J. Phys. Chem. A.* **2005**, *109*, 11752–11764.

[127] Wiberg, K. B.; Wang, Y.; Murphy, M. J.; Vaccaro, P. H. Temperature dependence of optical rotation: α-pinene, β-pinene, cis–pinane,ecamphene, camphor, and fenchone. *The J. Phys. Chem. A* **2004**, *108*, 5559–5563.

[128] Wiberg, K. B.; Wang, Y. G.; Wilson, S. M.; Vaccaro, P. H.; Cheeseman, J. R. Chiroptical properties of 2-chloropropionitrile. *J. Phys. Chem. A* **2005**, *109*, 3448–3453.

[129] Haghdani, S.; Hoff, B. H.; Koch, H.; Åstrand, P. O. Solvent Effects on Optical Rotation: On the Balance between Hydrogen Bonding and Shifts in Dihedral Angles. *J. Phys. Chem. A* **2017**, *121*, 4765–4777.

[130] Lipparini, F.; Egidi, F.; Cappelli, C.; Barone, V. The optical rotation of methyloxirane in aqueous solution: A never ending story? *J. Chem. Theory Comput.* **2013**, *9*, 1880–1884.

[131] Miertuš, S.; Scrocco, E.; Tomasi, J. Electrostatic interaction of a solute with a continuum. A direct utilizaion of AB initio molecular potentials for the prevision of solvent effects. *Chem. Phys.* **1981**, *55*, 117.

[132] Cancès, E.; Mennucci, B.; Tomasi, J. A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. *J. Chem. Phys.* **1997**, *107*, 3032–3041.

[133] Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chemical Reviews* **2005**, *105*, 2999–3093.

[134] Mennucci, B.; Cancès, E.; Tomasi, J. Evaluation of Solvent Effects in Isotropic and Anisotropic Dielectrics and in Ionic Solutions with a Unified Integral Equation Method: Theoretical Bases, Computational Implementation, and Numerical Applications. *The J. Phys. Chem. B* **1997**, *101*, 10506–10517.

[135] Lahiri, P.; Wiberg, K. B.; Vaccaro, P. H.; Caricato, M.; Crawford, T. D. Large solvation effect in the optical rotatory dispersion of norbornenone. *Angew. Chem. Int. Ed.* **2014**, *53*, 1386–1389.

[136] Lahiri, P.; Wiberg, K. B.; Vaccaro, P. H. Dispersive Optical Activity of (R)-Methylene Norbornene: Intrinsic Response and Solvation Effects. *J. Phys. Chem. A* **2017**, *5*, 8251–8266.

[137] Koch, H.; Jørgensen, P. Coupled cluster response functions. *J. Chem. Phys.* **1990**, *93*, 3333–3344.

[138] Pedersen, T. B.; Koch, H. Coupled cluster response functions revisited. *J. Chem. Phys.* **1997**, *106*, 8059–8072.

[139] Olsen, J.; Jørgensen, P. Linear and nonlinear response functions for an exact state and for an MCSCF state. *J. Chem. Phys.* **1985**, *82*, 3235–3264.

[140] Srebro, M.; Govind, N.; de Jong, W. A.; Autschbach, J. Optical Rotation Calculated with Time-Dependent Density Functional Theory: The OR45 Benchmark. *J. Phys. Chem. A* **2011**, *115*, 10930–10949.

[141] Čížek, J. In *Advances in Chemical Physics: Correlation Effects in Atoms and Molecules*, volume 14 ed.; LeFebvre, R., Moser, C., Eds.; John Wiley and Sons Inc.: New York, 1969; pp 35–89.

[142] Purvis, G. D.; Bartlett, R. J. A full coupled-cluster singles and doubles model: The inclusion of disconnected triples. *J. Chem. Phys.* **1982**, *76*, 1910–1918.

[143] Scuseria, G. E.; Janssen, C. L.; Schaefer, H. F. An efficient reformulation of the closed-shell coupled cluster single and double excitation (CCSD) equations. *J. Chem. Phys.* **1988**, *89*, 7382–7387.

[144] Scuseria, G. E.; Schaefer, H. F. Is coupled cluster singles and doubles (CCSD) more computationally intensive than quadratic configuration interaction (QCISD)? *J. Chem. Phys.* **1989**, *90*, 3700–3703.

[145] Lipparini, F.; Scalmani, G.; Mennucci, B.; Cancès, E.; Caricato, M.; Frisch, M. J. A variational formulation of the polarizable continuum model. *J. Chem. Phys.* **2010**, *133*, 014106.

[146] Amovilli, C.; Mennucci, B. Self-Consistent-Field Calculation of Pauli Repulsion and Dispersion Contributions to the Solvation Free Energy in the Polarizable Continuum Model. *The J. Phys. Chem. B* **1997**, *101*, 1051–1057.

[147] Autschbach, J.; Patchkovskii, S.; Ziegler, T.; van Gisbergen, S. J. A.; Jan Baerends, E. Chiroptical properties from time-dependent density functional theory. II. Optical rotations of small to medium sized organic molecules. *J. Chem. Phys.* **2002**, *117*, 581–592.

[148] Pulay, P. Localizability of dynamic electron correlation. *Chem. Phys. Lett.* **1983**, *100*, 151–154.

[149] Saebø, S.; Pulay, P. Local treatment of electron correlation. *Chem. Phys. Lett.* **1993**, *4*, 213–236.

[150] Russ, N. J.; Crawford, T. D. Local correlation in coupled cluster calculations of molecular response properties. *Chem. Phys. Lett.* **2004**, *400*, 104–111.

[151] Autschbach, J. Time-dependent density functional theory for calculating origin-independent optical rotation and rotatory strength tensors. *ChemPhysChem* **2011**, *12*, 3224–3235.

[152] Pople, J.; Krishnan, R.; Schlegel, B.; Binkley, J. Derivative Studies in Hartree-Fock and Moller-Plesset Theories. *Int. J. Quantum Chem.* **1979**, *16*, 225–241.

[153] Frisch, M.; Head-Gordon, M.; Pople, J. Direct analytic SCF second derivatives and electric field properties. *Chemical Physics* **1990**, *141*, 189–196.

[154] Stephens, P. J.; Devlin, F. J.; Cheeseman, J. R.; Frisch, M. J. Calculation of optical rotation using density functional theory. *J. Phys. Chem. A* **2001**, *105*, 5356–5371.

[155] Russ, N. J.; Crawford, T. D.; Tschumper, G. S. Real versus artifactual symmetry-breaking effects in Hartree-Fock, density-functional, and coupled-cluster methods. *J. Chem. Phys.* **2004**, *120*, 7298–7306.

[156] Howard, J. C.; Sowndarya, S. V.; Ansari, I. M.; Mach, T. J.; Baranowska-ŁAczkowska, A.; Crawford, T. D. Performance of Property-Optimized Basis Sets for Optical Rotation with Coupled Cluster Theory. *J. Phys. Chem. A* **2018**, *122*, 5962–5969.

[157] Martin, R. H. The Helicenes. *Angew. Chem. Int. Ed.* **1974**, *13*, 649–660.

[158] Shen, Y.; Chen, C. F. Helicenes: Synthesis and applications. *Chemical Reviews* **2012**, *112*, 1463–1535.

[159] Nuckolls, C.; Katz, T. J.; Castellanos, L. Aggregation of conjugated helical molecules. *J. Am. Chem. Soc.* **1996**, *118*, 3767–3768.

[160] Li, C.; Shi, J.; Xu, L.; Wang, Y.; Cheng, Y.; Wang, H. Syntheses and crystal structures of fused thiophenes: [7]Helicene and double helicene, a D2-symmetric dimer of 3,3′-bis(dithieno[2,3- b:3′,2′-d]thiophene). *J. Org. Chem.* **2009**, *74*, 408–411.

[161] Zak, J. K.; Miyasaka, M.; Rajca, S.; Lapkowski, M.; Rajca, A. Radical Cation of Helical , Cross-Conjugated β-Oligothiophene. *J. Am. Chem. Soc.* **2010**, *132*, 3246–3247.

[162] Miyasaka, M.; Pink, M.; Olankitwanit, A.; Rajca, S.; Rajca, A. Band gap of carbon-sulfur [n]helicenes. *Org. Lett.* **2012**, *14*, 3076–3079.

[163] Miyasaka, M.; Pink, M.; Rajca, S.; Rajca, A. Noncovalent interactions in the asymmetric synthesis of rigid, conjugated helical structures. *Angew. Chem. Int. Ed.* **2009**, *48*, 5954–5957.

[164] Rajca, A.; Miyasaka, M.; Pink, M.; Wang, H.; Rajca, S. Helically annelated and cross-conjugated oligothiophenes: Asymmetric synthesis, resolution, and characterization of a carbon-sulfur [7]helicene. *J. Am. Chem. Soc.* **2004**, *126*, 15211–15222.

[165] Fujikawa, T.; Segawa, Y.; Itami, K. Synthesis and Structural Features of Quadruple Helicenes: Highly Distorted p Systems Enabled by Accumulation of Helical Repulsions. *J. Am. Chem. Soc.* **2016**, *138*, 3587–3595.

[166] Yamamoto, Y.; Sakai, H.; Yuasa, J.; Araki, Y.; Wada, T.; Sakanoue, T.; Takenobu, T.; Kawai, T.; Hasobe, T. Synthetic Control of the Excited-State Dynamics and Circularly Polarized Luminescence of Fluorescent "push-Pull" Tetrathia[9]helicenes. *Chem. Eur. J.* **2016**, *22*, 4263–4273.

[167] Norsten, T. B.; Peters, A.; McDonald, R.; Wang, M.; Branda, N. R. Reversible [7]-thiahelicene formation using a 1,2-dithienylcyclopentene photochrome [13]. *J. Am. Chem. Soc.* **2001**, *123*, 7447–7448.

[168] Rajca, A.; Pink, M.; Xiao, S.; Miyasaka, M.; Rajca, S.; Das, K.; Plessel, K. Functionalized thiophene-based [7]helicene: Chirooptical properties versus electron delocalization. *J. Org. Chem.* **2009**, *74*, 7504–7513.

[169] Rose-Munch, F.; Li, M.; Rose, E.; Daran, J. C.; Bossi, A.; Licandro, E.; Mussini, P. R. Tetrathia[7]helicene-based complexes of ferrocene and (ν 5- cyclohexadienyl)tricarbonylmanganese: Synthesis and electrochemical studies. *Organometallics* **2012**, *31*, 92–104.

[170] Míšek, J.; Teplý, F.; Stará, I. G.; Tichý, M.; Šaman, D.; Císařová, I.; Vojtíšek, P.; Starý, I. A straightforward route to helically chiral N-heteroaromatic compounds: Practical synthesis

of racemic 1,14-diaza[5]helicene and optically pure 1- and 2-aza[6]helicenes. *Angew. Chem. Int. Ed.* **2008**, *47*, 3188–3191.

[171] Graule, S.; Crassous, J.; Réau, R.; Rudolph, M.; Autschbach, J.; Vanthuyne, N.; Roussel, C. Metal-bis(helicene) assemblies incorporating π-conjugated phosphole-azahelicene ligands: Impacting chiroptical properties by metal variation. *J. Am. Chem. Soc.* **2009**, *131*, 3183–3185.

[172] Ueda, A.; Wasa, H.; Suzuki, S.; Okada, K.; Sato, K.; Takui, T.; Morita, Y. Chiral stable phenalenyl radical: Synthesis, electronic-spin structure, and optical properties of [4]helicene-structured diazaphenalenyl. *Angew. Chem. Int. Ed.* **2012**, *51*, 6691–6695.

[173] Nakamura, K.; Furumi, S.; Takeuchi, M.; Shibuya, T.; Tanaka, K. Enantioselective synthesis and enhanced circularly polarized luminescence of S-shaped double azahelicenes. *J. Am. Chem. Soc.* **2014**, *136*, 5555–5558.

[174] Caronna, T.; Mele, A.; Famulari, A.; Mendola, D.; Fontana, F.; Juza, M.; Kamuf, M.; Zawatzky, K.; Trapp, O. A Combined Experimental and Theoretical Study on the Stereodynamics of Monoaza[5]helicenes: Solvent-Induced Increase of the Enantiomerization Barrier in 1-Aza-[5]helicene. *Chem. Eur. J.* **2015**, *21*, 13919–13924.

[175] Graule, S.; Rudolph, M.; Shen, W.; Gareth Williams, J. A.; Lescop, C.; Autschbach, J.; Crassous, J.; Réau, R. Assembly of π conjugated phosphole azahelicene derivatives into chiral coordination complexes: An experimental and theoretical study. *Chem. Eur. J.* **2010**, *16*, 5976–6005.

[176] Mendola, D.; Saleh, N.; Vanthuyne, N.; Roussel, C.; Toupet, L.; Castiglione, F.; Caronna, T.; Mele, A.; Crassous, J. Aza[6]helicene platinum complexes: Chirality control of cis-trans isomerism. *Angew. Chem. Int. Ed.* **2014**, *53*, 5786–5790.

[177] Mendola, D.; Saleh, N.; Hellou, N.; Vanthuyne, N.; Roussel, C.; Toupet, L.; Castiglione, F.; Melone, F.; Caronna, T.; Fontana, F.; MartÍ-Rujas, J.; Parisini, E.; Malpezzi, L.; Mele, A.;

Crassous, J. Synthesis and Structural Properties of Aza[n]helicene Platinum Complexes: Control of Cis and Trans Stereochemistry. *Inorg. Chem.* **2016**, *55*, 2009–2017.

[178] Ooyama, Y.; Ito, G.; Kushimoto, K.; Komaguchi, K.; Imae, I.; Harima, Y. Synthesis and fluorescence and electrochemical properties of D-π-A structural isomers of benzofuro[2,3-c]oxazolo[4,5-a]carbazole-type and benzofuro[2,3-c]oxazolo[5,4-a]carbazole-type fluorescent dyes. *Org. Biomol. Chem.* **2010**, *8*, 2756–2770.

[179] Shi, L.; Liu, Z.; Dong, G.; Duan, L.; Qiu, Y.; Jia, J.; Guo, W.; Zhao, D.; Cui, D.; Tao, X. Synthesis, structure, properties, and application of a carbazole-based diaza[7]helicene in a deep-blue-emitting OLED. *Chem. Eur. J.* **2012**, *18*, 8092–8099.

[180] Goto, K.; Yamaguchi, R.; Hiroto, S.; Ueno, H.; Kawai, T.; Shinokubo, H. Intermolecular oxidative annulation of 2-aminoanthracenes to diazaacenes and aza[7]helicenes. *Angew. Chem. Int. Ed.* **2012**, *51*, 10333–10336.

[181] Kaseyama, T.; Furumi, S.; Zhang, X.; Tanaka, K.; Takeuchi, M. Hierarchical assembly of a phthalhydrazide-functionalized helicene. *Angew. Chem. Int. Ed.* **2011**, *50*, 3684–3687.

[182] Žádný, J.; Jančařík, A.; Andronova, A.; Šámal, M.; Vacekchocholoušová, J.; Vacek, J.; Pohl, R.; Šaman, D.; Císařová, I.; Stará, I. G.; Starý, I. A general approach to optically pure [5]-, [6]-, and [7]heterohelicenes. *Angew. Chem. Int. Ed.* **2012**, *51*, 5857–5861.

[183] Guin, J.; Besnard, C.; Lacour, J. Synthesis, resolution, and stabilities of a cationic chromenoxanthene [4]helicene. *Org. Lett.* **2010**, *12*, 1748–1751.

[184] Herse, C.; Bas, D.; Krebs, F. C.; Bürgi, T.; Weber, J.; Wesolowski, T.; Laursen, B. W.; Lacour, J. A highly configurationally stable [4]heterohelicenium cation. *Angew. Chem. Int. Ed.* **2003**, *42*, 3162–3166.

[185] Kel, O.; Sherin, P.; Mehanna, N.; Laleu, B.; Lacour, J.; Vauthey, E. Excited-state properties of chiral [4]helicene cations. *Photochem. Photobiol. Sci.* **2012**, *11*, 623–631.

[186] Adriaenssens, L.; Severa, L.; Šálová, T.; Císařová, I.; Pohl, R.; Šaman, D.; Rocha, S. V.; Finney, N. S.; Pospíšil, L.; Slavíček, P.; Teplý, F. Helquats: A facile, modular, scalable route to novel helical dications. *Chem. Eur. J.* **2009**, *15*, 1072–1076.

[187] Pospíšil, L.; Bednárová, L.; Štěpánek, P.; Slavíček, P.; Vávra, J.; Hromadová, M.; Dlouhá, H.; Tarábek, J.; Teplý, F. Intense chiroptical switching in a dicationic helicene-like derivative: Exploration of a viologen-type redox manifold of a non-racemic helquat. *J. Am. Chem. Soc.* **2014**, *136*, 10826–10829.

[188] Coe, B. J.; Rusanova, D.; Joshi, V. D.; Sánchez, S.; Vávra, J.; Khobragade, D.; Severa, L.; Císařová, I.; Šaman, D.; Pohl, R.; Clays, K.; Depotter, G.; Brunschwig, B. S.; Teplý, F. Helquat Dyes: Helicene-like Push-Pull Systems with Large Second-Order Nonlinear Optical Responses. *J. Org. Chem.* **2016**, *81*, 1912–1920.

[189] Sakamaki, D.; Kumano, D.; Yashima, E.; Seki, S. A facile and versatile approach to double N-heterohelicenes: Tandem oxidative C-N couplings of N-heteroacenes via cruciform Dimers. *Angew. Chem. Int. Ed.* **2015**, *54*, 5404–5407.

[190] Sakamaki, D.; Kumano, D.; Yashima, E.; Seki, S. A double hetero[4]helicene composed of two phenothiazines: Synthesis, structural properties, and cationic states. *ChemComm* **2015**, *51*, 17237–17240.

[191] Nakano, K.; Oyama, H.; Nishimura, Y.; Nakasako, S.; Nozaki, K. λ 5-Phospha[7]helicenes: Synthesis, properties, and columnar aggregation with one-way chirality. *Angew. Chem. Int. Ed.* **2012**, *51*, 695–699.

[192] Yavari, K.; Moussa, S.; Ben Hassine, B.; Retailleau, P.; Voituriez, A.; Marinetti, A. 1H-phosphindoles as structural units in the synthesis of chiral helicenes. *Angew. Chem. Int. Ed.* **2012**, *51*, 6748–6752.

[193] Yavari, K.; Retailleau, P.; Voituriez, A.; Marinetti, A. Heterohelicenes with embedded P-

chiral 1H-phosphindole or dibenzophosphole units: Diastereoselective photochemical synthesis and structural characterization. *Chem. Eur. J.* **2013**, *19*, 9939–9947.

[194] Shyam Sundar, M.; Bedekar, A. V. Synthesis and Study of 7,12,17-Trioxa[11]helicene. *Org. Lett.* **2015**, *17*, 5808–5811.

[195] Oyama, H.; Nakano, K.; Harada, T.; Kuroda, R.; Naito, M.; Nobusawa, K.; Nozaki, K. Facile synthetic route to highly luminescent sila[7]helicene. *Org. Lett.* **2013**, *15*, 2104–2107.

[196] Katz, T. J.; Sudhakar, A.; Teasley, M. F.; Gilbert, A. M.; Geiger, W. E.; Robben, M. P.; Wuensch, M.; Ward, M. D. Synthesis and Properties of Optically Active Helical Metallocene Oligomers. *J. Am. Chem. Soc.* **1993**, *115*, 3182–3198.

[197] Anger, E.; Srebro, M.; Vanthuyne, N.; Toupet, L.; Rigaut, S.; Roussel, C.; Autschbach, J.; Crassous, J.; Réau, R. Ruthenium-vinylhelicenes: Remote metal-based enhancement and redox switching of the chiroptical properties of a helicene core. *J. Am. Chem. Soc.* **2012**, *134*, 15628–15631.

[198] Biet, T.; Fihey, A.; Cauchy, T.; Vanthuyne, N.; Roussel, C.; Crassous, J.; Avarvari, N. Ethylenedithio-tetrathiafulvalene-helicenes: Electroactive helical precursors with switchable chiroptical properties. *Chem. Eur. J.* **2013**, *19*, 13160–13167.

[199] Schweinfurth, D.; Zalibera, M.; Kathan, M.; Shen, C.; Mazzolini, M.; Trapp, N.; Crassous, J.; Gescheidt, G.; Diederich, F. Helicene Quinones: Redox-Triggered Chiroptical Switching and Chiral Recognition of the Semiquinone Radical Anion Lithium Salt by Electron Nuclear Double Resonance Spectroscopy. *J. Am. Chem. Soc.* **2014**, *136*, 13045–13052.

[200] Sapir, M.; Donckt, E. V. Intersystem crossing in the helicenes. *Chem. Phys. Lett.* **1975**, *36*, 108–110.

[201] Yamamoto, Y.; Sakai, H.; Yuasa, J.; Araki, Y.; Wada, T.; Sakanoue, T.; Takenobu, T.; Kawai, T.; Hasobe, T. Controlled Excited-State Dynamics and Enhanced Fluorescence Prop-

erty of Tetrasulfone[9]helicene by a Simple Synthetic Process. *J. Phys. Chem. C* **2016**, *120*, 7421–7427.

[202] Neto, B. A.; Lapis, A. A.; Da Silva Júnior, E. N.; Dupont, J. 2,1,3-benzothiadiazole and derivatives: Synthesis, properties, reactions, and applications in light technology of small molecules. *Eur. J. Org. Chem.* **2013**, *2*, 228–255.

[203] Neto, B. A.; Carvalho, P. H.; Correa, J. R. Benzothiadiazole Derivatives as Fluorescence Imaging Probes: Beyond Classical Scaffolds. *Acc. Chem. Res.* **2015**, *48*, 1560–1569.

[204] Chen, C. T. Evolution of red organic light-emitting diodes: Materials and devices. *Chem. Mater.* **2004**, *16*, 4389–4400.

[205] Justin Thomas, K. R.; Lin, J. T.; Velusamy, M.; Tao, Y.-T.; Chuen, C.-H. Color Tuning in Benzo[1,2,5]thiadiazole-Based Small Molecules by Amino Conjugation/Deconjugation: Bright Red-Light-Emitting Diodes. *Adv. Funct. Mater.* **2004**, *14*, 83–90.

[206] Zaumseil, J.; Donley, C. L.; Kim, J. S.; Friend, R. H.; Sirringhaus, H. Efficient top-gate, ambipolar, light-emitting field-effect transistors based on a green-light-emitting polyfluorene. *Adv. Mater.* **2006**, *18*, 2708–2712.

[207] Kono, T.; Kumaki, D.; Nishida, J.-i.; Sakanoue, T.; Kakita, M.; Tada, H.; Tokito, S.; Yamashita, Y. High-Performance and Light-Emitting n-Type Organic Field-Effect Transistors Based on Dithienylbenzothiadiazole and Related Heterocycles. *Chem. Mater.* **2007**, *19*, 1218–120.

[208] Zhang, M.; Tsao, H. N.; Pisula, W.; Yang, C.; Mishra, A. K.; Müllen, K. Field-effect transistors based on a benzothiadiazole-cyclopentadithiophene copolymer. *J. Am. Chem. Soc.* **2007**, *129*, 3472–3473.

[209] Yamashita, Y. Development of High-performance n-Type Organic Field-effect Transistors Based on Nitrogen Heterocycles. *Chem. Lett.* **2009**, *38*, 870–875.

[210] Velusamy, M.; Thomas, K. R.; Lin, J. T.; Hsu, Y. C.; Ho, K. C. Organic dyes incorporating low-band-gap chromophores for dye-sensitized solar cells. *Org. Lett.* **2005**, *7*, 1899–1902.

[211] Tang, Z. M.; Lei, T.; Jiang, K. J.; Song, Y. L.; Pei, J. Benzothiadiazole containing D-π-A conjugated compounds for dye-sensitized solar cells: Synthesis, properties, and photovoltaic performances. *Chem. Asian J.* **2010**, *5*, 1911–1917.

[212] Pop, F.; Avarvari, N. Covalent non-fused tetrathiafulvalene-acceptor systems. *ChemComm* **2016**, *52*, 7906–7927.

[213] Pop, F.; Amacher, A.; Avarvari, N.; Ding, J.; Daku, L. M. L.; Hauser, A.; Koch, M.; Hauser, J.; Liu, S. X.; Decurtins, S. Tetrathiafulvalene-benzothiadiazoles as redox-tunable donor-acceptor systems: Synthesis and photophysical study. *Chem. Eur. J.* **2013**, *19*, 2504–2514.

[214] Pop, F.; Seifert, S.; Hankache, J.; Ding, J.; Hauser, A.; Avarvari, N. Modulation of the charge transfer and photophysical properties in non-fused tetrathiafulvalene-benzothiadiazole derivatives. *Org. Biomol. Chem.* **2015**, *13*, 1040–1047.

[215] Martin, K. Hélicènes photo- et redox actifs. Molécules uniques et auto-assemblage sur surface. Ph.D. Dissertation, Universite d'Angers, 2019.

[216] Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396.

[217] Pina, J.; De Melo, J. S.; Breusov, D.; Scherf, U. Donor-acceptor-donor thienyl/bithienyl-benzothiadiazole/quinoxaline model oligomers: Experimental and theoretical studies. *Phys. Chem. Chem. Phys.* **2013**, *15*, 15204–15213.

[218] Schmidt, K.; Brovelli, S.; Coropceanu, V.; Beljonne, D.; Cornil, J.; Bazzini, C.; Caronna, T.;

Tubino, R.; Meinardi, F.; Shuai, Z.; Brédas, J. L. Intersystem crossing processes in nonplanar aromatic heterocyclic molecules. *J. Phys. Chem. A* **2007**, *111*, 10490–10499.

[219] Aharon, T.; Lemler, P.; Vaccaro, P. H.; Caricato, M. Comparison of measured and predicted specific optical rotation in gas and solution phases: A test for the polarizable continuum model of solvation. *Chirality* **2018**, *30*, 383–395.

Appendix A

Appendix for A Molecular Orbital Selection Approach for Fast Calculations of Specific Rotation with Density Functional Theory

Table A.1: Specific rotation (deg dm^{-1} (g/mL)$^{-1}$) computed with the aug-cc-pVDZ basis set and the AGn selection criterion. Full refers to the reference calculation with all molecular orbitals.

	CAM-B3LYP				B3LYP			
Molecule	Full	$n=5$	$n=4$	$n=3$	Full	$n=5$	$n=4$	$n=3$
1	-18.65	-19.50	-27.82	-4.72	-17.13	-17.98	-22.73	-7.41
2	70.86	71.90	73.80	76.42	76.00	76.73	79.62	86.73
3	71.76	72.35	67.71	54.59	54.61	55.01	52.47	63.23
4	169.19	169.37	166.91	140.24	165.25	165.33	164.68	168.82
5	57.04	57.19	54.65	32.90	58.50	58.65	58.70	48.30
6	15.31	14.71	12.95	46.14	11.77	11.37	8.53	34.52
7	86.58	85.92	82.58	103.99	102.12	101.52	100.86	125.48
8	-38.42	-39.04	-41.92	-41.11	-24.97	-25.56	-25.68	-30.28
9	-23.53	-24.83	-25.92	-33.98	-10.03	-10.76	-11.49	-20.24
10	67.88	67.55	67.93	66.91	80.78	80.56	82.37	70.94
11	-136.99	-137.09	-136.78	-143.03	-134.75	-134.90	-134.07	-137.41
12	-24.52	-24.92	-26.73	-22.92	-13.36	-13.53	-13.62	-7.24
13	-86.89	-86.96	-85.83	-94.40	-83.11	-83.13	-81.23	-91.68

14	-160.21	-160.62	-160.91	-168.33	-169.40	-169.88	-169.17	-177.19
15	-48.53	-47.95	-47.95	-37.56	-65.78	-65.22	-63.90	-49.72
16	-38.86	-39.06	-40.12	-43.29	-41.32	-41.56	-42.10	-45.97
17	126.14	125.82	132.42	124.63	131.22	131.18	135.77	134.10
18	2.32	2.31	2.29	5.69	1.89	1.88	1.54	6.20
19	9.95	9.69	7.84	3.42	8.46	8.17	7.12	2.82
20	-923.43	-922.18	-914.96	-860.93	-1192.48	-1191.78	-1180.66	-1129.04
21	42.24	42.17	44.51	38.84	58.18	58.32	59.08	54.56
22	-57.31	-57.15	-57.30	-89.72	-67.11	-67.21	-66.33	-78.39
23	85.82	85.91	87.55	84.00	94.08	94.50	95.10	98.18
24	95.97	95.42	95.84	93.15	116.21	115.82	116.02	113.14
25	80.36	80.65	84.12	79.86	84.02	84.30	87.65	87.76
26	68.21	67.40	68.46	72.91	79.69	78.90	79.47	82.21
27	44.72	44.78	45.11	58.16	43.04	43.24	43.47	47.66
28	28.48	28.83	30.45	31.92	29.63	29.75	30.01	34.24
29	128.21	128.37	128.53	124.11	149.84	149.99	150.26	147.61
30	94.34	94.32	96.41	90.34	111.35	111.38	112.06	103.92
31	58.93	58.42	58.62	61.14	45.71	45.23	45.20	52.81
32	6.68	6.81	6.94	-7.40	26.44	26.78	27.22	18.84
33	-212.58	-211.76	-210.48	-205.37	-254.35	-253.67	-252.12	-247.65
34	-9.52	-9.67	-9.00	-11.96	-8.97	-9.04	-10.22	-8.29
35	388.21	387.06	381.27	344.29	407.89	407.04	404.29	347.36
36	-226.91	-226.49	-225.65	-220.35	-347.72	-347.58	-345.59	-342.57
37	97.83	98.60	96.74	105.07	90.67	90.95	90.21	92.78
38	-2387.07	-2382.97	-2382.44	-2320.68	-2913.11	-2910.71	-2910.18	-2887.47
39	-3629.51	-3622.68	-3613.20	-3568.01	-4822.08	-4817.72	-4810.71	-4780.12
40	4567.17	4557.09	4547.70	4423.17	6364.12	6357.20	6349.74	6248.84

41	-3540.40	-3532.63	-3523.93	-3453.64	-4822.80	-4818.00	-4813.66	-4765.13
42	-4926.46	-4917.50	-4908.49	-4814.98	-7000.30	-6993.93	-6991.31	-6928.89
43	90.25	90.23	95.61	112.23	136.37	136.40	139.86	161.36
44	156.37	156.53	156.81	152.60	175.88	175.98	176.33	172.72
45	-104.57	-103.71	-104.96	-101.28	-96.04	-95.34	-95.63	-98.31
46	-176.19	-175.82	-177.15	-176.73	-214.64	-214.44	-213.53	-213.43
47	-118.95	-118.97	-119.62	-113.62	-145.99	-145.98	-145.45	-144.19
48	-39.09	-38.72	-40.03	-36.83	-51.76	-51.58	-52.56	-47.75
49	201.15	200.32	198.37	201.70	238.03	237.67	235.58	238.62
50	235.89	235.76	236.48	248.51	281.02	280.70	283.19	265.90
51	-274.59	-273.94	-272.87	-263.15	-328.47	-327.87	-326.89	-315.46

Table A.2: Specific rotation (deg dm^{-1} $(g/mL)^{-1}$) computed with the aug-cc-pVDZ basis set and the RG*n* selection criterion. Full refers to the reference calculation with all molecular orbitals.

	CAM-B3LYP				B3LYP			
Molecule	Full	$n=5$	$n=4$	$n=3$	Full	$n=5$	$n=4$	$n=3$
1	-18.65	-18.68	-18.72	-21.65	-17.13	-17.16	-17.71	-21.91
2	70.86	71.13	71.10	73.47	76.00	75.97	77.03	82.71
3	71.76	71.70	71.99	71.72	54.61	54.60	54.70	45.98
4	169.19	169.20	169.65	171.48	165.25	164.89	165.54	175.09
5	57.04	57.19	56.88	55.81	58.50	58.51	58.78	56.35
6	15.31	15.18	14.17	11.35	11.77	11.62	11.01	20.01
7	86.58	86.76	86.98	88.74	102.12	102.43	101.71	99.80
8	-38.42	-38.69	-38.89	-37.61	-24.97	-25.55	-25.43	-26.68
9	-23.53	-23.92	-24.47	-29.94	-10.03	-10.93	-11.49	-21.80
10	67.88	68.11	67.37	74.22	80.78	80.57	79.93	87.01
11	-136.99	-137.20	-136.98	-135.72	-134.75	-134.90	-133.80	-137.41
12	-24.52	-24.60	-24.75	-26.57	-13.36	-13.48	-13.34	-13.52
13	-86.89	-86.94	-86.96	-85.83	-83.11	-83.23	-83.51	-79.85
14	-160.21	-160.10	-160.54	-158.92	-169.40	-169.97	-169.99	-174.77
15	-48.53	-48.50	-47.95	-47.95	-65.78	-65.76	-64.61	-57.35
16	-38.86	-38.86	-39.58	-49.31	-41.32	-41.56	-42.14	-45.97
17	126.14	126.01	125.61	127.76	131.22	131.48	131.53	137.94
18	2.32	2.32	2.03	1.67	1.89	1.88	2.16	4.58
19	9.95	9.90	10.08	11.73	8.46	8.36	8.90	-5.19
20	-923.43	-923.12	-918.54	-877.29	-1192.48	-1189.45	-1159.24	-1125.22
21	42.24	42.28	42.43	39.54	58.18	58.37	61.57	44.76
22	-57.31	-57.33	-58.06	-55.11	-67.11	-67.34	-63.28	-94.81
23	85.82	85.81	86.02	87.65	94.08	94.07	94.36	93.69

24	95.97	95.89	95.39	93.88	116.21	116.15	115.53	115.96
25	80.36	80.27	80.35	80.89	84.02	83.93	84.26	89.21
26	68.21	68.27	67.87	69.68	79.69	79.69	78.90	78.92
27	44.72	44.72	44.70	44.10	43.04	43.01	43.34	46.27
28	28.48	28.45	28.64	28.73	29.63	29.69	29.77	32.44
29	128.21	128.21	128.34	129.33	149.84	149.88	150.17	147.34
30	94.34	94.32	94.43	95.02	111.35	111.45	111.38	112.85
31	58.93	58.87	57.75	58.10	45.71	45.09	43.74	52.02
32	6.68	6.71	7.00	7.51	26.44	26.45	26.71	26.90
33	-212.58	-211.93	-210.59	-197.62	-254.35	-254.10	-246.26	-279.05
34	-9.52	-9.53	-8.90	-8.91	-8.97	-8.13	-7.11	-10.01
35	388.21	387.89	385.53	380.48	407.89	407.04	404.16	350.90
36	-226.91	-227.08	-225.60	-229.23	-347.72	-344.99	-350.85	-310.13
37	97.83	98.45	96.53	114.73	90.67	90.83	91.69	108.40
38	-2387.07	-2381.89	-2338.30	-2301.54	-2913.11	-2897.12	-2882.08	-2796.70
39	-3629.51	-3614.15	-3569.90	-3473.17	-4822.08	-4781.33	-4730.69	-4872.17
40	4567.17	4543.45	4390.72	4164.11	6364.12	6275.65	6130.67	6303.69
41	-3540.40	-3528.61	-3487.45	-3318.14	-4822.80	-4773.72	-4720.40	-5197.12
42	-4926.46	-4900.12	-4767.51	-4514.49	-7000.30	-6916.30	-6850.35	-7472.16
43	90.25	90.09	90.19	102.12	136.37	136.24	137.57	153.09
44	156.37	156.43	156.49	156.57	175.88	175.93	175.96	176.10
45	-104.57	-104.31	-103.61	-97.33	-96.04	-95.46	-95.74	-97.46
46	-176.19	-175.79	-175.54	-181.53	-214.64	-214.56	-215.12	-202.86
47	-118.95	-118.96	-119.07	-119.63	-145.99	-145.94	-145.41	-142.32
48	-39.09	-38.74	-39.18	-40.36	-51.76	-51.58	-52.93	-43.88
49	201.15	200.85	199.96	198.00	238.03	237.55	235.66	231.55
50	235.89	235.82	235.37	236.82	281.02	280.26	279.75	280.39

51	-274.59	-274.64	-274.35	-274.08	-328.47	-328.14	-327.49	-314.14

Table A.3: Specific rotation (deg dm^{-1} $(g/mL)^{-1}$) computed with the aug-cc-pVDZ basis set and the ASn selection criterion. Full refers to the reference calculation with all molecular orbitals.

	CAM-B3LYP				B3LYP			
Molecule	Full	$n=5$	$n=4$	$n=3$	Full	$n=5$	$n=4$	$n=3$
1	-18.65	-18.67	-19.85	-50.43	-17.13	-17.19	-20.01	-51.45
2	70.86	71.13	72.87	83.71	76.00	76.13	78.09	85.57
3	71.76	71.84	71.17	60.57	54.61	54.65	54.57	49.40
4	169.19	169.32	169.83	169.09	165.25	165.29	165.40	166.29
5	57.04	57.26	57.59	50.52	58.50	58.60	58.95	49.60
6	15.31	15.21	14.05	16.22	11.77	11.67	10.73	10.81
7	86.58	86.58	88.43	96.33	102.12	102.05	103.01	114.49
8	-38.42	-38.29	-39.57	-33.68	-24.97	-24.84	-25.65	-23.32
9	-23.53	-23.54	-25.32	-29.93	-10.03	-9.98	-11.82	-18.17
10	67.88	68.12	67.35	72.08	80.78	81.03	80.27	80.51
11	-136.99	-137.11	-136.83	-137.84	-134.75	-134.84	-134.65	-130.85
12	-24.52	-24.77	-24.14	-19.51	-13.36	-13.56	-12.33	-10.99
13	-86.89	-86.97	-87.49	-82.93	-83.11	-83.14	-83.53	-88.35
14	-160.21	-160.29	-160.15	-160.85	-169.40	-169.44	-169.96	-177.14
15	-48.53	-47.97	-47.79	-37.39	-65.78	-64.94	-64.74	-50.71
16	-38.86	-38.80	-38.49	-40.22	-41.32	-41.31	-41.09	-54.64
17	126.14	126.17	125.57	129.66	131.22	131.28	130.76	135.18
18	2.32	2.32	2.11	-0.74	1.89	1.88	1.69	5.45
19	9.95	10.00	10.64	17.19	8.46	8.52	8.80	6.36
20	-923.43	-923.27	-918.01	-888.44	-1192.48	-1192.37	-1184.76	-1155.21
21	42.24	42.28	42.12	40.51	58.18	58.22	58.12	53.64
22	-57.31	-57.36	-59.44	-62.54	-67.11	-67.18	-67.91	-66.18
23	85.82	85.79	85.48	87.80	94.08	94.07	94.12	102.23

24	95.97	95.85	96.82	88.03	116.21	116.12	115.25	113.42
25	80.36	80.22	82.33	84.10	84.02	83.91	84.66	85.71
26	68.21	68.21	68.89	62.48	79.69	79.67	80.06	83.38
27	44.72	44.65	46.18	47.95	43.04	43.01	44.60	45.62
28	28.48	28.54	28.96	26.62	29.63	29.71	29.85	26.72
29	128.21	128.21	128.74	118.43	149.84	149.88	149.90	149.56
30	94.34	94.42	95.05	98.60	111.35	111.45	111.00	119.97
31	58.93	58.83	58.16	59.98	45.71	45.59	44.49	47.39
32	6.68	6.71	7.17	8.72	26.44	26.35	26.64	14.64
33	-212.58	-212.68	-211.39	-218.04	-254.35	-254.36	-252.70	-256.83
34	-9.52	-9.53	-10.42	-19.13	-8.97	-8.95	-8.21	-18.53
35	388.21	388.33	385.01	374.31	407.89	407.65	405.20	400.58
36	-226.91	-227.43	-226.78	-216.47	-347.72	-348.43	-347.42	-342.28
37	97.83	98.28	98.38	117.91	90.67	90.77	90.83	102.21
38	-2387.07	-2386.37	-2383.06	-2349.45	-2913.11	-2912.68	-2911.16	-2880.21
39	-3629.51	-3625.47	-3620.50	-3568.22	-4822.08	-4820.25	-4815.75	-4787.07
40	4567.17	4558.93	4550.09	4393.36	6364.12	6358.35	6351.85	6230.95
41	-3540.40	-3533.10	-3528.71	-3451.12	-4822.80	-4817.27	-4813.50	-4745.83
42	-4926.46	-4920.46	-4912.84	-4830.89	-7000.30	-6998.53	-6990.70	-6930.33
43	90.25	90.09	89.47	104.72	136.37	136.24	136.30	152.34
44	156.37	156.42	157.02	151.48	175.88	175.93	176.20	173.44
45	-104.57	-104.27	-103.42	-94.15	-96.04	-95.72	-95.96	-97.67
46	-176.19	-175.80	-174.92	-176.06	-214.64	-214.38	-213.81	-216.49
47	-118.95	-118.90	-116.89	-115.54	-145.99	-145.93	-144.56	-144.13
48	-39.09	-38.74	-39.08	-35.20	-51.76	-51.59	-51.81	-47.99
49	201.15	200.96	200.69	201.07	238.03	238.02	237.85	232.71
50	235.89	235.11	234.90	247.43	281.02	280.55	279.42	286.50

51	-274.59	-274.58	-273.41	-278.23	-328.47	-328.47	-326.57	-332.28

Table A.4: Specific rotation (deg dm^{-1} $(g/mL)^{-1}$) computed with the aug-cc-pVDZ basis set and the RSn selection criterion. Full refers to the reference calculation with all molecular orbitals.

	CAM-B3LYP				B3LYP			
Molecule	Full	$n=5$	$n=4$	$n=3$	Full	$n=5$	$n=4$	$n=3$
1	-18.65	-18.60	-18.64	-19.66	-17.13	-17.16	-17.14	-17.82
2	70.86	70.88	70.88	72.69	76.00	76.00	76.08	77.65
3	71.76	71.76	71.80	72.15	54.61	54.61	54.65	54.57
4	169.19	169.22	169.21	169.62	165.25	165.28	165.47	166.56
5	57.04	57.04	57.16	57.07	58.50	58.50	58.60	58.52
6	15.31	15.31	15.41	13.89	11.77	11.77	11.72	10.73
7	86.58	86.39	86.39	87.92	102.12	101.98	102.05	102.96
8	-38.42	-38.43	-38.27	-39.08	-24.97	-24.94	-24.84	-25.80
9	-23.53	-23.48	-23.51	-25.04	-10.03	-9.97	-9.97	-11.18
10	67.88	68.02	68.12	67.37	80.78	80.92	81.03	80.32
11	-136.99	-136.93	-136.95	-137.19	-134.75	-134.72	-134.84	-134.76
12	-24.52	-24.69	-24.67	-24.87	-13.36	-13.50	-13.48	-13.89
13	-86.89	-87.00	-86.96	-87.69	-83.11	-83.13	-83.13	-83.37
14	-160.21	-160.20	-160.10	-159.72	-169.40	-169.35	-169.82	-168.35
15	-48.53	-48.50	-48.50	-47.40	-65.78	-65.76	-65.59	-64.34
16	-38.86	-38.90	-38.76	-39.13	-41.32	-41.27	-41.31	-41.37
17	126.14	126.14	125.61	130.85	131.22	131.21	130.46	135.91
18	2.32	2.32	2.34	2.21	1.89	1.91	1.88	1.69
19	9.95	9.83	9.93	10.17	8.46	8.46	8.49	8.84
20	-923.43	-923.57	-922.72	-912.62	-1192.48	-1192.54	-1184.95	-1156.58
21	42.24	42.27	42.28	42.08	58.18	58.19	58.23	58.02
22	-57.31	-57.30	-57.29	-57.13	-67.11	-67.12	-67.11	-68.08
23	85.82	85.78	85.78	86.03	94.08	94.06	94.07	94.31

24	95.97	95.99	95.85	95.61	116.21	116.22	116.13	115.72
25	80.36	80.31	80.25	80.38	84.02	84.00	83.94	84.62
26	68.21	68.21	68.18	68.07	79.69	79.71	79.66	79.41
27	44.72	44.73	44.65	45.21	43.04	43.07	43.01	43.52
28	28.48	28.53	28.47	29.01	29.63	29.66	29.69	29.95
29	128.21	128.19	128.19	128.87	149.84	149.87	149.87	150.12
30	94.34	94.39	94.39	94.37	111.35	111.42	111.42	111.46
31	58.93	58.88	58.90	57.88	45.71	45.70	45.77	44.89
32	6.68	6.57	6.58	7.08	26.44	26.37	26.30	26.43
33	-212.58	-212.63	-212.26	-207.94	-254.35	-254.33	-253.17	-249.36
34	-9.52	-9.54	-9.55	-9.02	-8.97	-8.92	-8.95	-8.21
35	388.21	388.21	387.88	384.44	407.89	407.89	407.50	403.71
36	-226.91	-227.37	-227.17	-233.64	-347.72	-348.34	-347.70	-344.74
37	97.83	98.24	98.27	98.56	90.67	90.68	90.87	89.97
38	-2387.07	-2386.45	-2383.28	-2332.46	-2913.11	-2910.53	-2906.28	-2853.35
39	-3629.51	-3625.63	-3612.49	-3463.09	-4822.08	-4817.18	-4807.40	-4737.12
40	4567.17	4557.44	4540.72	4276.84	6364.12	6355.85	6307.32	6123.54
41	-3540.40	-3532.67	-3527.94	-3409.06	-4822.80	-4816.97	-4806.33	-4707.34
42	-4926.46	-4917.00	-4902.29	-4702.12	-7000.30	-6992.48	-6967.86	-6803.19
43	90.25	90.09	90.09	91.29	136.37	136.24	136.24	137.45
44	156.37	156.31	156.42	156.53	175.88	175.88	175.93	176.03
45	-104.57	-104.32	-104.40	-104.16	-96.04	-95.96	-95.59	-95.58
46	-176.19	-175.61	-175.80	-174.50	-214.64	-214.31	-213.71	-216.25
47	-118.95	-118.87	-118.87	-118.76	-145.99	-145.91	-145.96	-144.24
48	-39.09	-38.72	-38.83	-38.80	-51.76	-51.59	-51.55	-51.60
49	201.15	200.85	200.83	200.88	238.03	237.89	237.98	235.80
50	235.89	235.79	235.82	235.32	281.02	280.90	280.55	279.42

51	-274.59	-274.66	-274.58	-273.56	-328.47	-328.55	-328.51	-328.14

Table A.5: Specific rotation (deg dm^{-1} (g/mL)$^{-1}$) computed with the aug-cc-pVTZ basis set and the AGn selection criterion. Full refers to the reference calculation with all molecular orbitals.

	CAM-B3LYP				B3LYP			
Molecule	Full	$n=5$	$n=4$	$n=3$	Full	$n=5$	$n=4$	$n=3$
1	-11.28	-12.12	-18.19	-25.67	-9.61	-10.20	-15.35	-28.02
2	58.62	59.12	65.71	80.46	62.77	63.04	70.65	87.72
3	46.51	47.12	48.59	53.42	28.41	29.03	30.30	30.44
4	140.91	140.79	141.37	128.41	135.50	135.06	133.40	115.71
5	52.20	51.51	43.46	27.30	52.73	52.21	46.35	39.58
6	1.31	0.90	-1.70	49.75	-3.73	-4.16	-3.99	-5.65
7	79.66	79.71	78.52	83.74	93.39	93.53	92.85	127.10
8	-42.75	-43.27	-44.36	-31.62	-30.49	-30.94	-31.37	-23.59
9	-24.75	-25.97	-26.98	-35.49	-12.22	-13.39	-12.67	-16.25
10	65.50	65.02	66.36	86.42	78.50	78.15	77.89	86.10
11	-139.10	-138.83	-138.69	-123.30	-137.67	-137.47	-136.54	-131.62
12	-21.97	-23.01	-26.26	-8.14	-11.34	-11.92	-14.89	-0.11
13	-86.86	-87.59	-87.62	-106.20	-84.00	-84.21	-85.51	-92.62
14	-153.80	-153.24	-158.18	-141.09	-161.34	-161.10	-164.81	-145.47
15	-46.76	-44.84	-41.50	-46.67	-62.12	-60.42	-56.99	-55.13
16	-36.84	-36.77	-36.82	-30.14	-39.23	-39.24	-39.42	-41.56
17	129.68	129.15	127.47	118.75	134.52	134.06	132.91	124.20
18	1.87	1.78	1.72	4.29	1.53	1.24	1.88	4.57
19	9.49	9.76	14.66	-6.51	8.77	9.26	10.42	12.77
20	-909.64	-906.45	-890.12	-839.73	-1172.15	-1167.99	-1149.32	-1095.27
21	42.32	42.55	39.48	35.26	58.11	57.89	57.50	57.36
22	-54.46	-53.95	-54.66	-69.66	-63.73	-62.90	-64.17	-77.90
23	82.67	82.70	81.03	89.45	90.70	90.54	90.31	96.07

24	93.10	92.22	91.92	89.44	112.60	111.86	110.81	108.86
25	79.58	79.33	79.71	90.90	83.63	83.11	84.97	82.82
26	65.79	65.82	66.10	79.31	76.72	75.81	75.77	75.73
27	44.03	44.15	48.07	57.71	42.95	42.91	46.17	39.09
28	27.69	27.96	26.99	26.92	28.94	28.49	28.53	22.50
29	124.23	123.59	125.95	106.37	144.67	144.35	146.00	136.27
30	91.05	90.86	91.78	79.04	106.94	106.98	105.26	95.25
31	58.59	57.65	55.60	66.84	45.22	44.47	42.98	55.03
32	4.34	4.67	-0.45	3.86	23.47	23.67	22.03	19.99
33	-211.62	-210.48	-205.61	-189.68	-252.29	-250.92	-246.76	-240.44
34	-9.94	-8.97	-10.22	-7.56	-9.38	-8.30	-8.32	-2.45
35	393.35	391.77	388.47	328.01	412.68	411.48	408.85	366.50
36	-224.57	-225.87	-221.30	-198.36	-343.56	-343.72	-342.33	-326.95
37	98.51	96.76	96.60	96.26	91.38	90.47	90.73	95.42
38	-2394.97	-2393.20	-2383.99	-2332.43	-2922.54	-2921.24	-2915.52	-2884.40
39	-3648.62	-3645.59	-3633.38	-3587.82	-4847.45	-4845.14	-4837.34	-4804.76
40	4594.76	4587.82	4567.90	4489.63	6402.94	6397.36	6378.57	6300.60
41	-3565.88	-3560.97	-3546.70	-3486.96	-4857.38	-4854.09	-4841.13	-4798.54
42	-4957.50	-4949.82	-4937.74	-4875.50	-7046.15	-7040.73	-7029.41	-6972.44
43	81.65	81.94	87.69	99.35	126.46	126.85	130.08	141.60
44	154.45	154.39	152.77	142.91	173.83	173.79	172.14	173.36
45	-106.39	-106.00	-103.61	-88.72	-98.53	-98.42	-96.47	-73.50
46	-171.43	-171.28	-168.29	-159.41	-209.62	-209.54	-208.39	-199.71
47	-118.84	-118.12	-115.34	-117.33	-145.70	-145.09	-141.98	-145.85
48	-40.25	-40.34	-37.20	-24.20	-53.73	-53.97	-51.77	-47.66
49	201.86	200.93	198.88	213.19	238.65	237.84	236.06	234.89
50	235.77	233.96	236.62	227.01	280.53	278.85	280.47	273.07

51	-271.79	-270.42	-272.00	-268.05	-324.86	-323.83	-324.61	-320.75

Table A.6: Specific rotation (deg dm^{-1} $(g/mL)^{-1}$) computed with the aug-cc-pVTZ basis set and the RG*n* selection criterion. Full refers to the reference calculation with all molecular orbitals.

	CAM-B3LYP				B3LYP			
Molecule	Full	$n=5$	$n=4$	$n=3$	Full	$n=5$	$n=4$	$n=3$
1	-11.28	-11.40	-12.06	-19.66	-9.61	-9.72	-10.28	-15.66
2	58.62	58.99	59.63	64.62	62.77	63.02	62.83	70.15
3	46.51	46.96	46.88	53.48	28.41	28.45	28.24	37.14
4	140.91	140.88	140.95	139.72	135.50	135.54	133.20	121.99
5	52.20	51.99	51.53	41.65	52.73	52.62	52.21	44.69
6	1.31	1.41	0.34	-2.07	-3.73	-3.78	-2.88	-16.94
7	79.66	79.90	79.71	78.52	93.39	93.65	93.28	92.06
8	-42.75	-42.75	-44.61	-43.50	-30.49	-31.01	-32.05	-23.59
9	-24.75	-24.87	-25.71	-28.61	-12.22	-13.32	-12.28	-10.62
10	65.50	65.18	64.30	65.11	78.50	78.24	78.15	87.22
11	-139.10	-138.69	-137.55	-138.05	-137.67	-137.42	-136.54	-127.91
12	-21.97	-22.18	-22.96	-24.28	-11.34	-11.56	-12.33	-15.63
13	-86.86	-86.71	-87.42	-88.05	-84.00	-84.02	-84.68	-86.99
14	-153.80	-153.39	-152.68	-162.78	-161.34	-161.56	-162.75	-170.53
15	-46.76	-46.46	-44.41	-41.50	-62.12	-61.68	-59.59	-60.75
16	-36.84	-36.54	-37.42	-32.84	-39.23	-39.32	-39.44	-43.50
17	129.68	129.38	129.83	128.20	134.52	134.39	134.67	129.23
18	1.87	1.82	2.53	3.27	1.53	1.63	1.52	3.32
19	9.49	9.12	9.55	21.11	8.77	8.86	16.25	0.49
20	-909.64	-906.82	-890.55	-870.39	-1172.15	-1156.87	-1134.72	-1113.54
21	42.32	42.18	42.72	41.24	58.11	58.66	59.20	46.10
22	-54.46	-54.32	-53.99	-53.65	-63.73	-63.26	-62.12	-84.23
23	82.67	82.66	82.68	82.00	90.70	90.68	90.74	89.97

24	93.10	92.78	92.58	91.53	112.60	112.37	111.75	110.81
25	79.58	79.52	79.49	80.88	83.63	83.48	83.51	85.16
26	65.79	65.64	65.32	65.42	76.72	76.41	75.84	75.61
27	44.03	43.96	43.80	45.29	42.95	42.87	43.36	46.47
28	27.69	27.62	27.69	26.54	28.94	28.89	28.95	28.47
29	124.23	124.18	124.15	123.94	144.67	144.68	144.30	148.85
30	91.05	91.13	91.31	90.49	106.94	107.00	106.98	105.98
31	58.59	58.27	57.53	53.84	45.22	44.42	43.31	64.37
32	4.34	4.16	4.65	1.99	23.47	23.69	23.68	21.16
33	-211.62	-210.57	-207.15	-190.84	-252.29	-250.38	-239.11	-236.04
34	-9.94	-9.54	-9.18	-7.78	-9.38	-8.12	-9.19	-18.12
35	393.35	393.07	390.33	358.54	412.68	411.39	408.85	366.50
36	-224.57	-225.96	-223.52	-213.82	-343.56	-342.27	-338.96	-291.12
37	98.51	96.79	96.60	94.47	91.38	90.67	95.53	86.03
38	-2394.97	-2386.61	-2334.79	-2286.69	-2922.54	-2902.48	-2900.72	-2893.65
39	-3648.62	-3632.31	-3580.70	-3611.70	-4847.45	-4812.30	-4805.53	-5107.72
40	4594.76	4565.08	4442.02	4364.16	6402.94	6315.41	6229.14	6409.25
41	-3565.88	-3548.45	-3491.79	-3484.82	-4857.38	-4804.95	-4788.60	-5182.78
42	-4957.50	-4929.62	-4875.50	-4976.44	-7046.15	-6966.52	-7070.75	-7900.62
43	81.65	81.58	81.46	85.70	126.46	126.62	127.27	139.88
44	154.45	154.55	154.55	153.68	173.83	173.94	173.82	173.00
45	-106.39	-106.23	-105.68	-99.02	-98.53	-98.48	-96.72	-91.49
46	-171.43	-171.37	-170.63	-164.82	-209.62	-209.73	-206.41	-196.53
47	-118.84	-118.54	-118.12	-114.54	-145.70	-145.24	-143.70	-139.78
48	-40.25	-40.06	-40.23	-34.41	-53.73	-53.97	-52.01	-47.66
49	201.86	201.58	199.94	204.36	238.65	237.66	235.11	237.09
50	235.77	234.20	237.45	224.62	280.53	281.06	274.86	286.85

51	-271.79	-270.76	-271.68	-267.63	-324.86	-324.85	-319.96	-310.29

Table A.7: Specific rotation (deg dm^{-1} $(g/mL)^{-1}$) computed with the aug-cc-pVTZ basis set and the ASn selection criterion. Full refers to the reference calculation with all molecular orbitals.

	CAM-B3LYP				B3LYP			
Molecule	Full	$n=5$	$n=4$	$n=3$	Full	$n=5$	$n=4$	$n=3$
1	-11.28	-11.62	-15.05	-4.54	-9.61	-9.85	-13.12	-30.12
2	58.62	58.69	60.05	82.73	62.77	62.84	63.96	79.31
3	46.51	46.45	46.51	32.79	28.41	28.46	29.77	24.19
4	140.91	140.74	143.11	136.50	135.50	135.38	135.97	136.99
5	52.20	51.92	50.69	31.61	52.73	52.59	52.14	31.77
6	1.31	1.26	-0.28	9.60	-3.73	-3.86	-4.73	-7.81
7	79.66	79.97	79.33	75.82	93.39	93.61	93.63	85.39
8	-42.75	-42.52	-44.13	-51.19	-30.49	-30.34	-31.55	-30.78
9	-24.75	-24.64	-26.41	-31.03	-12.22	-12.10	-15.33	-19.95
10	65.50	65.22	62.93	56.65	78.50	78.20	76.98	73.64
11	-139.10	-138.89	-138.02	-143.66	-137.67	-137.47	-137.18	-134.63
12	-21.97	-22.41	-22.10	2.58	-11.34	-11.65	-11.35	-9.76
13	-86.86	-86.81	-85.46	-82.25	-84.00	-83.97	-84.63	-86.37
14	-153.80	-153.60	-153.78	-154.16	-161.34	-161.20	-160.89	-159.35
15	-46.76	-46.31	-44.32	-40.84	-62.12	-61.80	-59.98	-58.98
16	-36.84	-36.80	-36.82	-33.53	-39.23	-39.20	-39.54	-35.38
17	129.68	129.73	128.51	133.88	134.52	134.61	133.89	140.28
18	1.87	1.92	1.60	2.40	1.53	1.57	1.01	1.15
19	9.49	9.68	14.91	-12.21	8.77	8.98	12.63	19.17
20	-909.64	-907.94	-893.26	-873.81	-1172.15	-1170.21	-1156.20	-1117.60
21	42.32	42.33	41.82	51.78	58.11	58.03	58.48	67.06
22	-54.46	-54.17	-54.83	-77.71	-63.73	-63.46	-64.76	-83.24
23	82.67	82.64	83.81	84.45	90.70	90.57	92.47	91.19

24	93.10	92.71	91.53	92.47	112.60	112.24	110.81	106.25
25	79.58	79.46	79.83	93.72	83.63	83.45	84.89	85.96
26	65.79	65.48	65.09	68.92	76.72	76.46	76.01	73.34
27	44.03	44.03	43.90	41.65	42.95	42.89	44.96	40.32
28	27.69	27.65	27.30	14.70	28.94	28.92	28.85	19.37
29	124.23	124.20	124.20	126.11	144.67	144.65	143.91	152.35
30	91.05	91.19	91.49	90.16	106.94	107.06	106.93	102.81
31	58.59	58.08	57.12	76.43	45.22	44.86	43.76	50.20
32	4.34	4.23	2.91	1.68	23.47	23.72	22.93	24.75
33	-211.62	-211.00	-205.76	-199.10	-252.29	-251.76	-247.42	-239.73
34	-9.94	-9.60	-11.54	-4.74	-9.38	-8.98	-10.52	-5.84
35	393.35	393.41	390.27	374.71	412.68	413.14	409.52	390.15
36	-224.57	-226.38	-223.75	-218.15	-343.56	-344.39	-342.27	-330.69
37	98.51	96.93	102.05	84.44	91.38	90.59	92.98	92.30
38	-2394.97	-2394.03	-2391.06	-2378.36	-2922.54	-2922.02	-2919.76	-2911.53
39	-3648.62	-3646.54	-3640.21	-3633.15	-4847.45	-4846.21	-4841.63	-4824.65
40	4594.76	4590.59	4576.34	4512.75	6402.94	6399.58	6387.99	6300.36
41	-3565.88	-3563.14	-3553.78	-3526.71	-4857.38	-4855.43	-4847.14	-4806.27
42	-4957.50	-4953.31	-4943.86	-4939.53	-7046.15	-7043.43	-7035.67	-6996.78
43	81.65	81.77	84.20	95.78	126.46	126.70	129.09	133.70
44	154.45	154.55	153.53	156.77	173.83	173.89	172.71	173.60
45	-106.39	-106.13	-105.50	-102.69	-98.53	-98.48	-97.55	-93.22
46	-171.43	-170.90	-171.07	-163.98	-209.62	-209.18	-209.83	-209.15
47	-118.84	-118.26	-117.33	-111.63	-145.70	-145.23	-144.62	-138.70
48	-40.25	-39.84	-39.10	-25.02	-53.73	-53.57	-53.04	-33.47
49	201.86	201.81	199.29	198.58	238.65	238.49	236.98	237.85
50	235.77	234.89	232.62	220.81	280.53	279.65	276.86	283.21

51	-271.79	-271.02	-269.88	-268.32	-324.86	-324.09	-320.40	-323.06

Table A.8: Specific rotation (deg dm^{-1} $(g/mL)^{-1}$) computed with the aug-cc-pVTZ basis set and the RSn selection criterion. Full refers to the reference calculation with all molecular orbitals.

	CAM-B3LYP				B3LYP			
Molecule	Full	$n=5$	$n=4$	$n=3$	Full	$n=5$	$n=4$	$n=3$
1	-11.28	-11.39	-11.52	-11.70	-9.61	-9.74	-9.82	-9.77
2	58.62	58.53	58.62	59.34	62.77	62.78	62.84	63.96
3	46.51	46.58	46.56	47.29	28.41	28.47	28.56	30.10
4	140.91	140.99	141.00	141.85	135.50	135.58	135.07	134.59
5	52.20	51.95	51.96	51.40	52.73	52.58	52.59	52.05
6	1.31	1.22	1.33	0.73	-3.73	-3.80	-3.75	-4.52
7	79.66	79.57	79.63	79.33	93.39	93.34	93.42	93.44
8	-42.75	-42.82	-42.60	-43.20	-30.49	-30.31	-30.33	-30.76
9	-24.75	-24.77	-24.57	-25.91	-12.22	-12.15	-12.06	-13.50
10	65.50	65.47	65.35	64.81	78.50	78.37	78.32	76.54
11	-139.10	-138.82	-138.67	-139.10	-137.67	-137.38	-137.20	-137.49
12	-21.97	-22.21	-22.06	-23.41	-11.34	-11.51	-11.52	-11.61
13	-86.86	-86.86	-86.78	-84.50	-84.00	-83.98	-83.94	-82.87
14	-153.80	-153.69	-153.57	-153.60	-161.34	-161.09	-161.26	-165.57
15	-46.76	-46.61	-46.51	-44.99	-62.12	-61.96	-61.78	-60.34
16	-36.84	-36.87	-36.76	-37.08	-39.23	-39.20	-39.19	-39.96
17	129.68	129.60	129.30	131.95	134.52	134.46	134.08	137.76
18	1.87	1.88	1.93	1.81	1.53	1.55	1.61	1.28
19	9.49	9.41	9.32	10.61	8.77	9.00	9.22	11.59
20	-909.64	-908.69	-906.88	-889.12	-1172.15	-1170.29	-1156.90	-1124.61
21	42.32	42.36	42.28	41.80	58.11	58.18	58.24	59.03
22	-54.46	-54.54	-54.43	-54.46	-63.73	-63.89	-63.79	-63.36
23	82.67	82.71	82.59	84.00	90.70	90.72	90.63	90.84

24	93.10	93.11	92.81	92.30	112.60	112.55	112.32	111.48
25	79.58	79.61	79.46	79.20	83.63	83.62	83.45	84.23
26	65.79	65.72	65.57	66.42	76.72	76.69	76.50	76.06
27	44.03	44.07	44.00	44.60	42.95	42.98	42.89	43.29
28	27.69	27.73	27.67	27.58	28.94	28.94	28.89	29.30
29	124.23	124.22	124.24	124.17	144.67	144.70	144.72	144.63
30	91.05	91.16	91.12	91.66	106.94	107.05	107.02	107.23
31	58.59	58.33	58.20	57.43	45.22	45.06	44.82	43.66
32	4.34	4.19	4.26	4.33	23.47	23.45	23.69	22.46
33	-211.62	-211.48	-210.33	-204.19	-252.29	-252.00	-250.77	-244.78
34	-9.94	-9.60	-9.65	-8.94	-9.38	-9.02	-9.03	-10.11
35	393.35	393.36	393.41	390.27	412.68	412.73	412.31	404.60
36	-224.57	-225.76	-225.39	-230.03	-343.56	-344.37	-341.71	-325.13
37	98.51	98.28	97.02	102.14	91.38	90.86	90.77	93.88
38	-2394.97	-2394.64	-2389.43	-2373.70	-2922.54	-2921.26	-2927.04	-2892.25
39	-3648.62	-3646.56	-3632.51	-3614.94	-4847.45	-4843.98	-4844.04	-4803.14
40	4594.76	4588.00	4571.66	4448.36	6402.94	6394.64	6355.31	6219.89
41	-3565.88	-3562.67	-3553.32	-3519.52	-4857.38	-4852.16	-4843.49	-4766.17
42	-4957.50	-4951.66	-4928.97	-4974.84	-7046.15	-7038.39	-7010.30	-6974.12
43	81.65	81.58	81.82	82.82	126.46	126.47	126.85	128.86
44	154.45	154.44	154.61	154.41	173.83	173.83	173.93	173.85
45	-106.39	-106.34	-106.32	-105.49	-98.53	-98.58	-98.38	-97.62
46	-171.43	-171.06	-170.92	-171.16	-209.62	-209.29	-209.34	-208.34
47	-118.84	-118.48	-118.48	-117.82	-145.70	-145.42	-145.17	-143.35
48	-40.25	-39.99	-40.11	-40.17	-53.73	-53.94	-53.70	-53.60
49	201.86	202.11	201.95	200.37	238.65	238.70	237.93	234.94
50	235.77	235.35	234.82	231.11	280.53	279.98	279.66	275.80

51	-271.79	-271.63	-271.15	-270.39	-324.86	-324.72	-324.11	-320.41

Table A.9: Number of basis functions selected for the AG*n* calculations with aug-cc-pVDZ (Number of occupied orbitals, number of virtual orbitals). Full refers to the reference calculation with all molecular orbitals.

	CAM-B3LYP				B3LYP		
Molecule	Full	$n=5$	$n=4$	$n=3$	$n=5$	$n=4$	$n=3$
1	16, 130	12, 113	11, 85	10, 51	12, 114	11, 86	10, 52
2	20, 167	15, 152	14, 116	14, 72	15, 155	14, 116	13, 74
3	20, 130	12, 116	11, 89	9, 55	12, 118	11, 90	10, 57
4	73, 559	51, 516	51, 422	49, 260	51, 520	51, 429	46, 261
5	20, 185	15, 164	14, 125	13, 80	15, 164	15, 126	13, 81
6	23, 205	17, 186	16, 146	14, 88	17, 190	16, 148	15, 94
7	20, 176	15, 157	14, 116	12, 73	15, 157	14, 117	12, 77
8	28, 186	18, 161	17, 124	15, 82	18, 163	17, 127	15, 82
9	37, 186	20, 164	17, 133	16, 82	20, 167	18, 134	17, 82
10	24, 149	15, 128	14, 96	12, 64	15, 130	14, 100	12, 68
11	24, 149	15, 130	14, 105	12, 67	15, 131	14, 108	12, 67
12	28, 232	21, 196	19, 145	16, 91	21, 198	19, 148	16, 98
13	32, 269	24, 222	22, 178	19, 107	24, 223	22, 181	19, 112
14	19, 168	14, 148	13, 111	12, 62	14, 148	14, 113	12, 69
15	20, 167	15, 143	14, 109	12, 67	15, 143	14, 112	12, 72
16	31, 188	20, 174	18, 135	16, 82	20, 174	19, 136	16, 85
17	19, 168	14, 153	14, 124	12, 77	14, 155	14, 126	12, 77
18	34, 84	15, 83	12, 67	9, 39	15, 83	11, 67	9, 40
19	30, 244	22, 217	22, 157	17, 106	22, 220	22, 159	18, 112
20	29, 227	21, 202	20, 148	16, 88	21, 207	19, 150	17, 95
21	42, 355	31, 297	29, 225	24, 139	31, 301	29, 230	26, 146
22	42, 355	31, 296	29, 228	25, 127	31, 300	29, 232	26, 143

23	31, 243	23, 206	21, 143	17, 86	23, 206	21, 144	18, 95
24	35, 280	26, 238	24, 172	19, 102	26, 243	24, 172	20, 107
25	35, 280	26, 239	23, 173	20, 105	26, 239	24, 176	20, 113
26	39, 317	29, 263	27, 192	21, 116	29, 267	27, 195	23, 120
27	39, 317	29, 261	27, 199	23, 117	29, 264	27, 203	24, 125
28	39, 317	29, 262	27, 195	21, 117	29, 269	27, 198	23, 126
29	31, 225	23, 191	19, 137	17, 83	23, 192	21, 137	17, 91
30	35, 262	26, 216	23, 162	19, 95	26, 220	23, 161	19, 98
31	38, 336	28, 281	27, 211	22, 128	28, 283	27, 215	23, 135
32	38, 336	28, 282	27, 210	24, 130	28, 285	27, 212	24, 133
33	41, 338	30, 293	30, 229	23, 143	30, 296	30, 230	24, 146
34	38, 318	28, 269	27, 208	22, 129	28, 272	26, 212	24, 136
35	31, 279	23, 253	23, 180	19, 114	23, 254	23, 187	19, 113
36	67, 532	48, 490	47, 378	43, 235	48, 495	47, 382	43, 242
37	74, 620	54, 562	53, 431	49, 303	54, 567	53, 435	50, 320
38	86, 656	60, 604	60, 485	57, 300	60, 605	60, 490	53, 304
39	92, 696	64, 629	64, 498	63, 301	64, 631	64, 502	58, 307
40	103, 666	65, 606	63, 477	61, 286	66, 608	63, 486	57, 291
41	99, 753	69, 692	69, 557	66, 335	69, 691	69, 553	61, 338
42	24, 167	15, 153	14, 121	13, 74	15, 153	14, 121	14, 76
43	27, 188	20, 181	19, 117	15, 85	20, 180	19, 118	15, 88
44	44, 317	33, 280	30, 196	25, 136	33, 285	32, 197	25, 145
45	38, 336	28, 289	27, 224	22, 136	28, 292	27, 228	23, 141
46	44, 358	32, 336	31, 249	27, 165	32, 337	32, 256	29, 163
47	46, 329	34, 308	31, 214	27, 154	34, 306	32, 215	27, 162
48	38, 336	28, 282	28, 218	22, 128	28, 282	27, 223	23, 140
49	38, 336	28, 286	27, 216	22, 136	28, 290	27, 224	23, 143

50	27, 224	20, 197	20, 149	15, 103	20, 199	20, 150	15, 115
51	27, 224	20, 196	20, 147	16, 99	20, 200	20, 148	16, 105

Table A.10: Number of basis functions selected for the RGn calculations with aug-cc-pVDZ (Number of occupied orbitals, number of virtual orbitals). Full refers to the reference calculation with all molecular orbitals.

	CAM-B3LYP				B3LYP		
Molecule	Full	$n=5$	$n=4$	$n=3$	$n=5$	$n=4$	$n=3$
1	16, 130	13, 130	12, 124	11, 92	13, 130	12, 117	11, 89
2	20, 167	17, 167	15, 158	14, 118	16, 166	15, 148	14, 110
3	20, 130	15, 128	12, 108	11, 71	12, 127	12, 102	11, 64
4	24, 167	17, 167	15, 142	14, 105	48, 295	44, 183	38, 81
5	20, 185	17, 184	15, 160	14, 124	15, 182	15, 153	14, 116
6	23, 205	19, 203	17, 178	16, 139	17, 200	17, 157	16, 104
7	20, 176	17, 176	15, 170	14, 128	15, 173	15, 147	14, 111
8	28, 186	20, 183	18, 149	17, 108	18, 164	17, 130	15, 84
9	37, 186	24, 184	18, 149	17, 102	20, 163	18, 134	16, 82
10	24, 149	17, 148	15, 120	14, 81	15, 144	15, 109	13, 72
11	24, 149	17, 145	15, 117	14, 77	15, 131	14, 109	12, 67
12	28, 232	22, 232	21, 194	19, 144	21, 232	21, 179	19, 133
13	32, 269	26, 269	24, 222	22, 178	25, 268	24, 211	22, 163
14	19, 168	14, 166	14, 140	13, 99	14, 156	14, 120	12, 74
15	20, 167	15, 167	15, 143	14, 109	15, 167	15, 132	14, 97
16	31, 188	20, 186	20, 148	17, 105	20, 174	19, 132	16, 85
17	19, 168	16, 166	14, 140	14, 111	14, 162	14, 134	13, 85
18	34, 84	20, 84	13, 75	10, 53	15, 82	11, 65	9, 39
19	30, 244	22, 243	22, 200	20, 146	22, 201	20, 145	17, 84
20	29, 227	21, 213	21, 160	16, 102	21, 168	18, 120	15, 61
21	42, 355	31, 353	31, 282	28, 210	31, 282	28, 209	23, 121
22	42, 355	31, 352	31, 274	28, 194	31, 278	29, 201	23, 114

23	31, 243	24, 242	23, 219	21, 155	23, 242	23, 203	21, 141
24	35, 280	27, 280	26, 248	24, 179	27, 279	26, 232	23, 168
25	35, 280	30, 279	26, 254	24, 190	27, 279	26, 227	23, 164
26	39, 317	32, 316	29, 282	27, 217	31, 316	29, 267	27, 194
27	39, 317	33, 316	29, 281	27, 210	29, 316	29, 249	25, 187
28	39, 317	32, 317	29, 286	27, 216	29, 316	29, 254	27, 189
29	31, 225	23, 224	23, 202	20, 146	23, 224	23, 181	19, 133
30	35, 262	30, 262	26, 255	24, 179	27, 261	26, 224	23, 166
31	38, 336	28, 330	28, 260	26, 173	28, 270	27, 191	21, 115
32	38, 336	30, 335	28, 288	27, 216	29, 333	28, 269	27, 183
33	41, 338	30, 298	30, 232	24, 147	30, 261	28, 185	22, 98
34	38, 318	28, 317	28, 249	25, 190	28, 249	26, 189	21, 108
35	31, 279	23, 279	23, 230	22, 158	23, 254	23, 188	19, 114
36	67, 532	48, 511	47, 395	43, 256	47, 411	44, 284	38, 169
37	74, 620	54, 530	53, 410	48, 262	54, 449	50, 338	45, 178
38	73, 559	51, 432	49, 270	43, 160	53, 325	50, 196	40, 65
39	86, 656	60, 483	57, 298	50, 175	58, 333	54, 203	37, 62
40	92, 696	64, 462	62, 285	52, 158	57, 300	52, 185	38, 47
41	103, 666	63, 499	61, 301	53, 183	61, 326	56, 189	32, 42
42	99, 753	68, 503	66, 308	55, 170	15, 160	15, 134	14, 85
43	27, 188	20, 188	20, 173	18, 113	20, 187	19, 137	16, 106
44	44, 317	34, 317	33, 298	32, 204	33, 316	33, 253	28, 183
45	38, 336	28, 333	28, 262	26, 187	28, 298	27, 233	23, 148
46	44, 358	32, 349	32, 271	29, 185	32, 315	31, 223	26, 142
47	46, 329	36, 329	34, 305	31, 213	34, 324	34, 244	28, 193
48	38, 336	29, 333	28, 267	27, 187	28, 282	27, 218	22, 137
49	38, 336	30, 333	28, 273	27, 197	28, 287	27, 215	22, 139

50	27, 224	20, 224	20, 185	18, 139	20, 184	19, 137	15, 86
51	27, 224	20, 223	20, 177	18, 130	20, 178	18, 133	15, 78

Table A.11: Number of basis functions selected for the ASn calculations with aug-cc-pVDZ (Number of occupied orbitals, number of virtual orbitals). Full refers to the reference calculation with all molecular orbitals.

	CAM-B3LYP				B3LYP		
Molecule	Full	$n=5$	$n=4$	$n=3$	$n=5$	$n=4$	$n=3$
1	16, 130	12, 128	12, 99	10, 64	12, 129	11, 99	10, 64
2	20, 167	17, 167	15, 131	14, 88	17, 167	15, 131	14, 90
3	20, 130	12, 128	12, 104	10, 65	12, 128	12, 105	10, 64
4	73, 559	63, 551	51, 463	46, 286	65, 552	51, 467	45, 286
5	20, 185	19, 183	15, 145	14, 95	19, 184	15, 147	14, 93
6	23, 205	19, 203	17, 166	16, 112	19, 203	17, 166	16, 111
7	20, 176	15, 175	15, 135	14, 84	15, 175	15, 138	12, 85
8	28, 186	18, 185	18, 146	15, 87	18, 185	18, 146	16, 88
9	37, 186	18, 184	18, 144	16, 91	18, 185	18, 146	16, 89
10	24, 149	15, 148	15, 120	12, 72	15, 148	15, 120	12, 74
11	24, 149	15, 147	15, 118	13, 71	15, 147	15, 117	13, 71
12	28, 232	21, 231	20, 171	16, 115	21, 231	20, 172	17, 113
13	32, 269	24, 268	24, 197	18, 127	24, 268	24, 197	18, 128
14	19, 168	14, 164	14, 130	12, 79	14, 165	14, 129	12, 78
15	20, 167	15, 164	15, 128	12, 77	15, 163	15, 129	12, 78
16	31, 188	20, 187	20, 153	16, 100	20, 187	20, 154	16, 102
17	19, 168	18, 167	14, 143	13, 102	18, 167	14, 144	13, 105
18	34, 84	13, 78	12, 72	9, 36	13, 78	12, 73	9, 38
19	30, 244	22, 243	21, 186	18, 118	22, 243	21, 187	18, 118
20	29, 227	21, 224	20, 176	18, 112	21, 224	20, 177	18, 112
21	42, 355	31, 351	30, 265	24, 159	31, 351	30, 266	24, 155
22	42, 355	31, 352	30, 263	23, 157	31, 352	30, 263	24, 163

23	31, 243	23, 242	23, 185	19, 114	23, 242	23, 185	19, 117
24	35, 280	26, 279	25, 208	21, 130	26, 279	26, 212	21, 130
25	35, 280	26, 279	25, 208	22, 137	26, 279	26, 209	21, 141
26	39, 317	29, 315	28, 230	24, 150	29, 315	28, 235	22, 145
27	39, 317	29, 316	28, 235	23, 149	29, 316	28, 235	24, 152
28	39, 317	29, 314	28, 231	23, 146	29, 315	28, 234	24, 146
29	31, 225	23, 224	22, 169	18, 115	23, 224	23, 171	19, 118
30	35, 262	26, 261	24, 195	19, 130	26, 261	26, 196	20, 131
31	38, 336	28, 330	28, 259	23, 149	28, 330	28, 258	22, 148
32	38, 336	28, 331	28, 257	23, 150	28, 333	28, 259	23, 150
33	41, 338	30, 331	29, 254	23, 155	30, 334	29, 255	23, 157
34	38, 318	28, 317	27, 238	23, 143	28, 317	27, 238	22, 148
35	31, 279	29, 279	23, 227	22, 142	28, 279	23, 228	23, 142
36	67, 532	48, 531	48, 423	41, 251	48, 530	48, 426	41, 247
37	74, 620	54, 613	53, 453	45, 271	54, 615	53, 454	45, 267
38	86, 656	74, 644	60, 541	53, 316	79, 644	60, 542	53, 316
39	92, 696	64, 663	64, 546	56, 300	64, 668	64, 550	56, 296
40	103, 666	63, 645	63, 519	55, 297	63, 646	63, 520	55, 288
41	99, 753	81, 736	69, 616	60, 347	81, 739	69, 619	60, 349
42	24, 167	17, 166	15, 138	14, 93	17, 166	15, 140	14, 94
43	27, 188	20, 188	20, 146	16, 98	20, 188	20, 148	16, 97
44	44, 317	33, 317	33, 234	26, 169	33, 317	33, 238	26, 169
45	38, 336	28, 331	28, 256	24, 147	28, 331	28, 257	24, 157
46	44, 358	32, 354	32, 275	26, 163	32, 354	32, 276	25, 161
47	46, 329	34, 328	33, 244	26, 159	34, 328	33, 243	26, 159
48	38, 336	28, 330	28, 259	23, 159	28, 331	28, 259	24, 157
49	38, 336	28, 332	28, 259	23, 156	28, 332	28, 259	23, 157

50	27, 224	20, 221	19, 169	16, 111	20, 221	19, 169	16, 111
51	27, 224	20, 224	19, 173	16, 108	20, 224	19, 175	16, 109

Table A.12: Number of basis functions selected for the RSn calculations with aug-cc-pVDZ (Number of occupied orbitals, number of virtual orbitals). Full refers to the reference calculation with all molecular orbitals.

	CAM-B3LYP				B3LYP		
Molecule	Full	$n=5$	$n=4$	$n=3$	$n=5$	$n=4$	$n=3$
1	16, 130	15, 130	12, 130	12, 110	13, 130	12, 130	12, 102
2	20, 167	19, 167	19, 167	15, 139	19, 167	17, 165	15, 128
3	20, 130	14, 130	12, 129	12, 111	12, 130	12, 128	12, 106
4	73, 559	63, 551	51, 458	45, 285	51, 511	51, 391	42, 212
5	20, 185	20, 185	19, 184	15, 156	20, 185	19, 184	15, 153
6	23, 205	23, 205	22, 205	17, 178	23, 205	19, 205	17, 169
7	20, 176	17, 176	17, 176	15, 143	17, 176	15, 175	15, 137
8	28, 186	20, 186	18, 186	18, 151	20, 186	18, 186	18, 148
9	37, 186	21, 186	18, 185	18, 149	22, 186	18, 186	18, 158
10	24, 149	15, 149	15, 148	15, 126	15, 149	15, 148	15, 127
11	24, 149	16, 149	15, 149	15, 129	15, 149	15, 147	15, 120
12	28, 232	22, 232	21, 232	21, 200	22, 232	21, 232	21, 192
13	32, 269	25, 269	24, 269	24, 223	24, 269	24, 269	24, 209
14	19, 168	17, 168	14, 166	14, 136	14, 166	14, 151	14, 115
15	20, 167	15, 167	15, 167	15, 147	15, 167	15, 166	15, 140
16	31, 188	25, 188	20, 188	20, 167	20, 188	20, 187	20, 158
17	19, 168	19, 168	16, 154	14, 122	18, 168	14, 148	14, 111
18	34, 84	19, 84	13, 79	13, 76	15, 84	13, 78	12, 73
19	30, 244	25, 244	22, 244	22, 215	22, 244	22, 242	21, 186
20	29, 227	21, 227	21, 208	20, 149	21, 225	20, 179	18, 114
21	42, 355	35, 355	31, 354	31, 296	31, 355	31, 354	30, 271
22	42, 355	31, 355	31, 354	31, 290	31, 355	31, 353	30, 272

23	31, 243	25, 243	23, 243	23, 202	25, 243	23, 243	23, 202
24	35, 280	30, 280	26, 280	26, 235	28, 280	26, 280	26, 230
25	35, 280	31, 280	26, 280	26, 246	30, 280	26, 280	26, 229
26	39, 317	32, 317	29, 317	29, 259	32, 317	29, 317	29, 262
27	39, 317	34, 317	29, 316	29, 259	33, 317	29, 316	29, 257
28	39, 317	31, 317	29, 317	29, 262	31, 317	29, 316	29, 259
29	31, 225	23, 225	23, 225	23, 199	23, 225	23, 225	23, 183
30	35, 262	26, 262	26, 262	26, 215	26, 262	26, 262	26, 217
31	38, 336	30, 336	28, 331	28, 260	30, 336	28, 333	28, 263
32	38, 336	28, 336	28, 334	28, 269	28, 336	28, 311	28, 235
33	41, 338	30, 338	30, 325	29, 234	30, 338	30, 304	29, 222
34	38, 318	31, 318	28, 318	28, 254	30, 318	28, 317	27, 238
35	31, 279	31, 279	25, 276	23, 214	31, 279	23, 259	23, 179
36	67, 532	50, 532	48, 483	44, 331	48, 527	48, 411	39, 238
37	74, 620	58, 620	54, 615	53, 471	54, 620	54, 610	53, 441
38	86, 656	68, 615	60, 491	51, 270	60, 586	57, 392	49, 206
39	92, 696	64, 634	63, 469	53, 239	64, 599	58, 386	52, 207
40	103, 666	63, 637	63, 507	55, 276	63, 588	60, 397	52, 219
41	99, 753	69, 707	68, 542	59, 290	69, 642	63, 384	54, 208
42	24, 167	19, 167	19, 166	15, 142	19, 167	17, 163	15, 130
43	27, 188	20, 188	20, 188	19, 137	20, 188	20, 188	18, 129
44	44, 317	38, 317	33, 317	33, 296	36, 317	33, 317	33, 283
45	38, 336	30, 336	28, 335	28, 274	28, 336	28, 328	28, 248
46	44, 358	33, 358	32, 354	32, 287	32, 355	32, 297	26, 182
47	46, 329	34, 329	34, 329	34, 278	34, 329	34, 326	32, 225
48	38, 336	32, 336	28, 336	28, 267	32, 336	28, 334	28, 264
49	38, 336	32, 336	28, 336	28, 271	28, 336	28, 315	27, 237

50	27, 224	24, 224	20, 224	20, 186	21, 224	20, 221	19, 169
51	27, 224	22, 224	20, 224	20, 179	22, 224	20, 223	19, 169

Table A.13: Number of basis functions selected for the AG*n* calculations with aug-cc-pVTZ (Number of occupied orbitals, number of virtual orbitals). Full refers to the reference calculation with all molecular orbitals.

	CAM-B3LYP				B3LYP		
Molecule	Full	$n=5$	$n=4$	$n=3$	$n=5$	$n=4$	$n=3$
1	16, 306	12, 166	11, 96	9, 63	12, 168	11, 98	10, 66
2	20, 394	15, 215	14, 136	12, 83	15, 218	14, 139	13, 88
3	20, 306	12, 163	11, 107	9, 65	12, 165	11, 108	10, 69
4	73, 1261	51, 732	51, 514	49, 309	51, 742	51, 524	47, 323
5	20, 440	15, 229	14, 150	12, 93	15, 231	15, 152	13, 93
6	23, 483	17, 262	16, 173	14, 110	17, 265	16, 178	15, 114
7	20, 417	15, 212	14, 136	11, 82	15, 216	14, 137	11, 87
8	28, 436	18, 230	17, 146	13, 93	18, 234	17, 147	14, 95
9	37, 436	22, 235	17, 155	14, 98	22, 237	18, 156	16, 103
10	24, 348	15, 188	14, 118	12, 72	15, 191	14, 119	12, 74
11	24, 348	15, 185	14, 117	11, 71	15, 186	14, 117	12, 73
12	28, 547	21, 291	19, 174	16, 106	21, 298	19, 174	16, 114
13	32, 635	24, 331	22, 203	19, 123	24, 338	22, 207	18, 125
14	19, 395	14, 199	13, 134	12, 80	14, 201	13, 134	12, 80
15	20, 394	15, 207	14, 125	12, 78	15, 212	14, 125	12, 81
16	31, 433	20, 255	18, 157	16, 99	20, 256	19, 159	16, 103
17	19, 395	14, 223	14, 155	11, 93	14, 226	14, 157	12, 96
18	34, 190	17, 128	11, 81	9, 53	17, 130	12, 83	9, 54
19	30, 568	22, 329	20, 193	17, 117	22, 333	20, 196	18, 122
20	29, 523	21, 305	20, 175	16, 104	21, 306	20, 178	17, 112
21	42, 832	31, 460	29, 271	24, 154	31, 461	29, 274	26, 165
22	42, 832	31, 435	29, 271	24, 161	31, 444	29, 275	26, 166

23	31, 567	23, 297	20, 174	17, 111	23, 302	20, 178	17, 118
24	35, 655	26, 330	24, 203	19, 122	26, 341	23, 208	19, 129
25	35, 655	26, 330	23, 208	20, 127	26, 336	23, 209	19, 130
26	39, 743	28, 369	27, 233	21, 138	29, 377	27, 234	22, 147
27	39, 743	29, 367	27, 231	21, 136	29, 368	27, 238	23, 142
28	39, 743	28, 361	26, 231	19, 134	29, 371	26, 235	24, 144
29	31, 521	23, 274	20, 158	17, 102	23, 282	20, 164	17, 106
30	35, 609	26, 309	22, 187	18, 111	26, 318	23, 189	19, 118
31	38, 790	28, 383	27, 255	23, 153	28, 386	27, 256	22, 160
32	38, 790	28, 385	26, 259	24, 150	28, 391	27, 263	24, 154
33	41, 787	30, 450	30, 261	25, 163	30, 455	30, 271	25, 167
34	38, 744	28, 421	26, 246	23, 151	28, 426	26, 249	22, 155
35	31, 659	23, 337	23, 221	18, 137	23, 343	23, 223	19, 139
36	67, 1221	48, 682	47, 467	43, 288	48, 687	46, 475	44, 303
37	74, 1444	54, 865	53, 540	47, 338	54, 880	53, 555	50, 346
38	86, 1478	60, 855	60, 593	60, 352	60, 864	60, 602	55, 360
39	92, 1564	64, 862	64, 604	64, 358	64, 872	64, 606	59, 369
40	103, 1497	68, 835	63, 563	62, 333	68, 848	63, 579	58, 340
41	99, 1695	69, 947	69, 661	69, 394	69, 965	69, 664	63, 406
42	24, 394	15, 215	14, 144	11, 89	15, 219	14, 144	12, 91
43	27, 433	20, 252	18, 156	15, 99	20, 252	18, 159	15, 103
44	44, 738	33, 420	29, 252	23, 152	33, 423	31, 259	25, 164
45	38, 790	28, 398	27, 269	22, 159	28, 401	27, 276	23, 167
46	44, 830	32, 456	31, 315	28, 181	32, 459	32, 313	29, 182
47	46, 759	34, 460	30, 277	27, 176	34, 462	31, 292	27, 183
48	38, 790	28, 394	27, 258	23, 161	28, 396	27, 266	23, 165
49	38, 790	28, 402	27, 259	23, 161	28, 410	27, 263	23, 169

50	27, 525	20, 313	19, 180	15, 115	20, 313	19, 182	15, 121
51	27, 525	20, 310	19, 174	15, 105	20, 313	19, 179	16, 115

Table A.14: Number of basis functions selected for the RG*n* calculations with aug-cc-pVTZ (Number of occupied orbitals, number of virtual orbitals). Full refers to the reference calculation with all molecular orbitals.

	CAM-B3LYP				B3LYP		
Molecule	Full	$n=5$	$n=4$	$n=3$	$n=5$	$n=4$	$n=3$
1	16, 306	13, 237	12, 179	11, 113	12, 232	12, 171	11, 106
2	20, 394	17, 310	15, 227	14, 148	15, 292	15, 210	14, 130
3	20, 306	12, 220	12, 150	11, 101	12, 195	11, 131	11, 87
4	73, 1261	51, 538	50, 327	42, 179	49, 375	44, 206	37, 75
5	20, 440	16, 332	15, 234	14, 155	15, 301	15, 210	14, 135
6	23, 483	18, 362	17, 258	16, 168	17, 297	16, 205	16, 134
7	20, 417	15, 306	15, 212	14, 136	15, 284	15, 198	14, 126
8	28, 436	21, 315	17, 209	17, 139	18, 237	17, 150	14, 95
9	37, 436	23, 314	19, 206	17, 134	22, 235	18, 153	16, 97
10	24, 348	17, 252	15, 166	14, 111	15, 208	15, 130	13, 80
11	24, 348	17, 236	15, 158	14, 107	15, 187	14, 117	12, 74
12	28, 547	22, 416	21, 308	19, 187	21, 378	21, 276	19, 167
13	32, 635	25, 472	24, 341	22, 209	24, 454	24, 314	22, 192
14	19, 395	14, 266	14, 190	13, 126	14, 218	14, 142	12, 91
15	20, 394	15, 279	15, 205	14, 125	15, 257	15, 184	13, 117
16	31, 433	20, 291	19, 214	17, 125	20, 258	19, 162	16, 104
17	19, 395	14, 293	14, 205	14, 139	14, 248	14, 183	14, 117
18	34, 190	21, 145	13, 96	9, 66	17, 124	10, 80	9, 51
19	30, 568	22, 367	22, 267	20, 163	22, 297	20, 173	17, 101
20	29, 523	21, 313	20, 187	16, 121	21, 216	18, 138	15, 67
21	42, 832	31, 547	31, 390	28, 248	31, 415	29, 257	23, 139
22	42, 832	31, 565	31, 389	29, 246	31, 382	29, 241	23, 135

23	31, 567	23, 456	23, 329	21, 200	23, 396	23, 289	19, 172
24	35, 655	28, 513	26, 374	24, 227	27, 471	26, 337	23, 206
25	35, 655	27, 504	26, 363	24, 231	26, 448	26, 311	23, 193
26	39, 743	32, 584	29, 434	27, 283	30, 529	29, 383	27, 235
27	39, 743	30, 542	29, 378	27, 245	29, 523	29, 352	27, 228
28	39, 743	31, 552	29, 391	26, 241	29, 524	29, 356	25, 231
29	31, 521	23, 425	23, 300	21, 198	23, 363	23, 271	20, 153
30	35, 609	29, 516	26, 361	24, 245	26, 436	26, 318	23, 188
31	38, 790	28, 528	28, 360	26, 227	28, 361	27, 232	21, 132
32	38, 790	29, 574	28, 395	27, 270	28, 533	28, 356	26, 239
33	41, 787	30, 457	30, 276	25, 167	30, 350	28, 213	22, 108
34	38, 744	28, 497	28, 363	25, 227	28, 381	26, 231	21, 124
35	31, 659	23, 402	23, 282	20, 164	23, 342	23, 223	19, 139
36	67, 1221	48, 754	47, 519	44, 331	48, 549	44, 361	38, 196
37	74, 1444	54, 864	53, 540	47, 334	53, 590	51, 367	45, 186
38	86, 1478	60, 587	58, 342	48, 186	56, 392	50, 216	37, 65
39	92, 1564	64, 566	61, 319	52, 170	59, 410	54, 226	31, 61
40	103, 1497	63, 602	62, 359	52, 200	58, 362	53, 196	32, 43
41	99, 1695	69, 593	68, 333	52, 176	63, 377	56, 206	23, 41
42	24, 394	18, 299	15, 208	14, 139	15, 250	15, 162	14, 114
43	27, 433	20, 314	20, 242	17, 142	20, 273	20, 206	16, 120
44	44, 738	34, 645	33, 445	32, 298	33, 507	33, 395	29, 226
45	38, 790	28, 527	28, 359	26, 240	28, 419	27, 287	24, 177
46	44, 830	32, 503	32, 354	29, 210	32, 419	30, 281	27, 159
47	46, 759	35, 609	34, 449	30, 266	34, 491	33, 380	28, 215
48	38, 790	28, 543	28, 368	27, 244	28, 396	27, 265	23, 165
49	38, 790	28, 507	28, 355	26, 228	28, 389	27, 251	22, 157

50	27, 525	20, 357	20, 274	18, 158	20, 282	19, 164	15, 96
51	27, 525	20, 357	20, 265	18, 158	20, 271	18, 161	15, 93

Table A.15: Number of basis functions selected for the ASn calculations with aug-cc-pVTZ (Number of occupied orbitals, number of virtual orbitals). Full refers to the reference calculation with all molecular orbitals.

	CAM-B3LYP				B3LYP		
Molecule	Full	$n=5$	$n=4$	$n=3$	$n=5$	$n=4$	$n=3$
1	16, 306	12, 216	11, 138	8, 76	12, 217	11, 138	10, 79
2	20, 394	19, 298	14, 193	14, 110	19, 299	14, 194	14, 112
3	20, 306	12, 214	12, 132	9, 75	12, 213	12, 132	9, 75
4	73, 1261	63, 969	51, 606	45, 347	65, 972	51, 606	45, 352
5	20, 440	19, 321	15, 210	14, 115	19, 323	15, 215	14, 117
6	23, 483	19, 358	17, 222	15, 131	19, 359	17, 226	16, 133
7	20, 417	15, 299	14, 184	12, 99	15, 300	14, 185	14, 100
8	28, 436	18, 314	17, 186	15, 107	18, 312	18, 183	16, 109
9	37, 436	18, 311	17, 190	15, 113	18, 312	17, 193	15, 114
10	24, 348	15, 249	14, 151	12, 86	15, 251	14, 153	11, 82
11	24, 348	15, 242	15, 154	12, 90	15, 245	15, 157	12, 88
12	28, 547	21, 368	20, 235	15, 123	21, 371	20, 236	16, 125
13	32, 635	24, 436	23, 273	18, 145	24, 438	23, 273	18, 143
14	19, 395	14, 274	14, 175	12, 100	14, 276	14, 175	12, 101
15	20, 394	15, 279	14, 183	13, 96	15, 280	14, 181	13, 96
16	31, 433	20, 332	18, 223	16, 116	20, 332	19, 225	16, 119
17	19, 395	18, 301	14, 204	13, 124	18, 302	14, 205	12, 122
18	34, 190	13, 138	12, 91	9, 53	13, 139	12, 92	9, 55
19	30, 568	22, 374	21, 238	17, 132	22, 379	21, 241	18, 132
20	29, 523	21, 367	20, 235	18, 130	21, 367	20, 236	18, 127
21	42, 832	31, 547	29, 332	24, 180	31, 548	29, 326	24, 173
22	42, 832	31, 545	30, 332	23, 180	31, 547	30, 328	24, 183

23	31, 567	23, 392	21, 241	19, 133	23, 398	21, 242	18, 130
24	35, 655	26, 455	24, 268	20, 149	26, 458	24, 267	20, 148
25	35, 655	26, 450	24, 267	21, 148	26, 452	24, 277	22, 151
26	39, 743	29, 506	27, 301	22, 167	29, 504	27, 302	23, 169
27	39, 743	29, 496	27, 294	23, 166	29, 500	28, 297	24, 170
28	39, 743	29, 504	27, 302	21, 161	29, 500	27, 303	21, 164
29	31, 521	23, 368	22, 235	18, 129	23, 369	22, 233	19, 130
30	35, 609	26, 422	25, 264	20, 142	26, 424	25, 265	20, 144
31	38, 790	28, 518	27, 317	23, 170	28, 520	28, 317	22, 168
32	38, 790	28, 516	27, 315	23, 171	28, 523	27, 320	22, 170
33	41, 787	30, 525	28, 322	23, 177	30, 529	28, 323	23, 179
34	38, 744	28, 498	26, 307	23, 158	28, 501	26, 302	23, 163
35	31, 659	25, 479	23, 291	20, 162	24, 483	23, 293	21, 163
36	67, 1221	50, 899	48, 553	41, 303	48, 901	47, 557	40, 311
37	74, 1444	54, 952	53, 552	45, 308	54, 956	53, 551	45, 310
38	86, 1478	71, 1130	60, 690	53, 377	73, 1136	60, 697	53, 390
39	92, 1564	64, 1116	64, 683	55, 347	64, 1120	64, 684	55, 350
40	103, 1497	63, 1079	63, 655	55, 332	63, 1087	63, 655	55, 337
41	99, 1695	77, 1265	69, 767	60, 404	79, 1274	69, 768	60, 401
42	24, 394	19, 303	15, 188	14, 112	19, 304	15, 189	14, 115
43	27, 433	20, 312	19, 218	16, 114	20, 313	19, 221	15, 115
44	44, 738	33, 516	32, 337	25, 179	33, 521	32, 340	26, 179
45	38, 790	28, 547	28, 324	23, 175	28, 546	27, 329	23, 178
46	44, 830	32, 568	32, 355	26, 185	32, 567	32, 354	25, 183
47	46, 759	34, 543	32, 342	26, 178	34, 530	32, 345	26, 178
48	38, 790	28, 537	28, 323	22, 174	28, 540	27, 321	22, 175
49	38, 790	28, 533	28, 330	22, 177	28, 537	28, 333	23, 179

50	27, 525	20, 360	19, 231	15, 132	20, 361	19, 234	16, 123
51	27, 525	20, 366	19, 227	16, 128	20, 369	19, 230	16, 128

Table A.16: Number of basis functions selected for the ASn calculations with aug-cc-pVTZ (Number of occupied orbitals, number of virtual orbitals). Full refers to the reference calculation with all molecular orbitals.

	CAM-B3LYP				B3LYP		
Molecule	Full	$n=5$	$n=4$	$n=3$	$n=5$	$n=4$	$n=3$
1	16, 306	13, 297	12, 249	12, 162	12, 293	12, 231	12, 153
2	20, 394	19, 383	19, 314	15, 211	19, 379	19, 299	14, 194
3	20, 306	14, 302	12, 253	12, 167	12, 291	12, 218	12, 134
4	73, 1261	61, 955	51, 591	45, 326	51, 762	48, 469	42, 234
5	20, 440	19, 433	19, 340	15, 232	19, 431	19, 339	15, 231
6	23, 483	22, 474	19, 374	17, 239	22, 475	19, 375	17, 244
7	20, 417	17, 410	17, 317	15, 220	17, 410	17, 307	15, 194
8	28, 436	20, 431	18, 353	18, 236	19, 431	18, 339	18, 221
9	37, 436	23, 431	18, 350	18, 225	20, 431	18, 338	18, 213
10	24, 348	15, 344	15, 269	15, 178	15, 340	15, 254	15, 154
11	24, 348	15, 344	15, 269	15, 177	15, 344	15, 268	15, 170
12	28, 547	21, 538	21, 441	21, 298	21, 536	21, 402	20, 260
13	32, 635	24, 624	24, 478	23, 309	24, 623	24, 472	23, 298
14	19, 395	17, 389	14, 318	14, 209	16, 334	14, 222	13, 135
15	20, 394	15, 380	15, 307	15, 203	15, 376	15, 284	14, 188
16	31, 433	22, 422	20, 365	20, 259	20, 415	20, 349	20, 234
17	19, 395	19, 334	14, 230	14, 142	18, 321	14, 218	13, 135
18	34, 190	13, 178	13, 159	12, 103	13, 176	13, 141	12, 93
19	30, 568	22, 554	22, 427	22, 293	22, 547	22, 387	21, 248
20	29, 523	21, 447	21, 308	20, 172	21, 378	20, 247	18, 134
21	42, 832	31, 813	31, 610	30, 388	31, 807	31, 578	30, 357
22	42, 832	31, 817	31, 632	31, 416	31, 809	31, 600	30, 377

23	31, 567	25, 552	23, 430	22, 276	25, 554	23, 458	23, 298
24	35, 655	30, 641	26, 520	26, 340	28, 638	26, 491	26, 315
25	35, 655	30, 640	26, 509	26, 329	30, 638	26, 493	25, 316
26	39, 743	31, 727	29, 567	28, 365	31, 727	29, 562	29, 362
27	39, 743	33, 729	29, 578	28, 369	33, 726	29, 564	29, 349
28	39, 743	31, 729	29, 581	29, 386	31, 727	29, 579	29, 381
29	31, 521	23, 509	23, 452	23, 292	23, 510	23, 418	23, 272
30	35, 609	26, 599	26, 532	26, 348	26, 595	26, 495	26, 317
31	38, 790	30, 779	28, 590	28, 378	28, 760	28, 526	28, 323
32	38, 790	28, 779	28, 598	28, 380	28, 653	28, 429	27, 256
33	41, 787	30, 723	30, 490	28, 297	30, 648	30, 446	27, 250
34	38, 744	30, 728	28, 560	27, 370	30, 719	28, 510	26, 313
35	31, 659	31, 645	25, 479	23, 291	31, 580	23, 387	23, 230
36	67, 1221	54, 1021	48, 652	43, 383	48, 874	47, 531	39, 286
37	74, 1444	58, 1409	54, 1018	53, 600	54, 1335	54, 884	52, 505
38	86, 1478	61, 993	59, 605	50, 307	60, 770	56, 463	49, 227
39	92, 1564	64, 887	61, 542	53, 263	64, 785	58, 465	52, 213
40	103, 1497	63, 1051	63, 634	55, 323	63, 769	57, 469	52, 227
41	99, 1695	69, 1073	68, 635	58, 324	69, 806	62, 454	54, 221
42	24, 394	19, 388	19, 309	15, 196	19, 367	15, 272	15, 169
43	27, 433	20, 416	20, 290	18, 201	20, 406	20, 273	18, 184
44	44, 738	36, 727	33, 685	33, 439	36, 727	33, 674	33, 435
45	38, 790	28, 777	28, 591	28, 367	28, 721	28, 502	27, 304
46	44, 830	32, 806	32, 573	32, 361	32, 594	32, 374	25, 203
47	46, 759	34, 746	34, 683	34, 454	34, 716	34, 489	30, 305
48	38, 790	33, 780	28, 597	28, 384	33, 774	28, 577	28, 354
49	38, 790	30, 776	28, 573	28, 359	28, 667	28, 451	27, 265

50	27, 525	22, 514	20, 410	19, 285	20, 498	20, 360	19, 226
51	27, 525	22, 509	20, 389	19, 266	22, 505	20, 357	19, 222

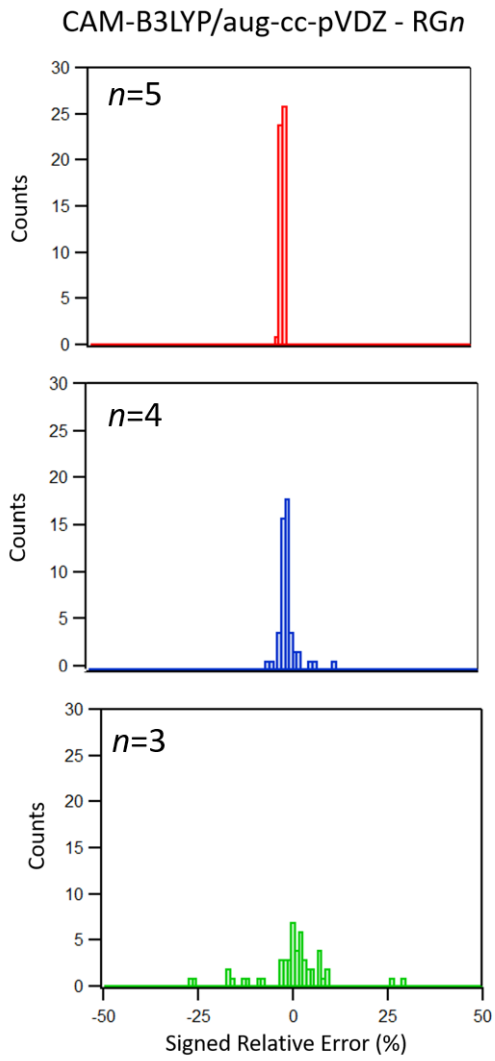


Figure A.1: Histogram of the signed relative errors for each of the 51 molecules. These results are calculated with CAM-B3LYP, using relative thresholds determined from $\langle m \rangle^2$.

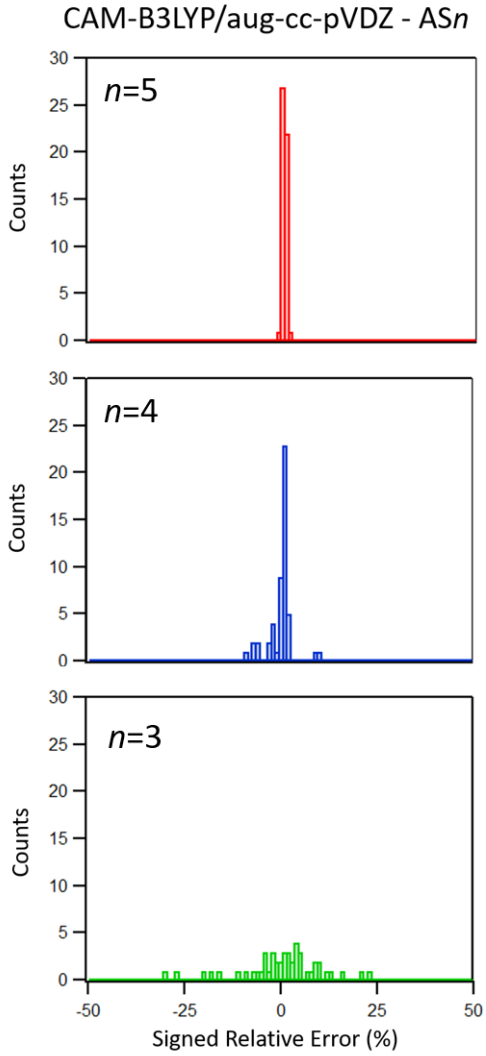


Figure A.2: Histogram of the signed relative errors for each of the 51 molecules. These results are calculated with CAM-B3LYP, using absolute thresholds determined from $\langle|\mu|\rangle\langle m\rangle$.

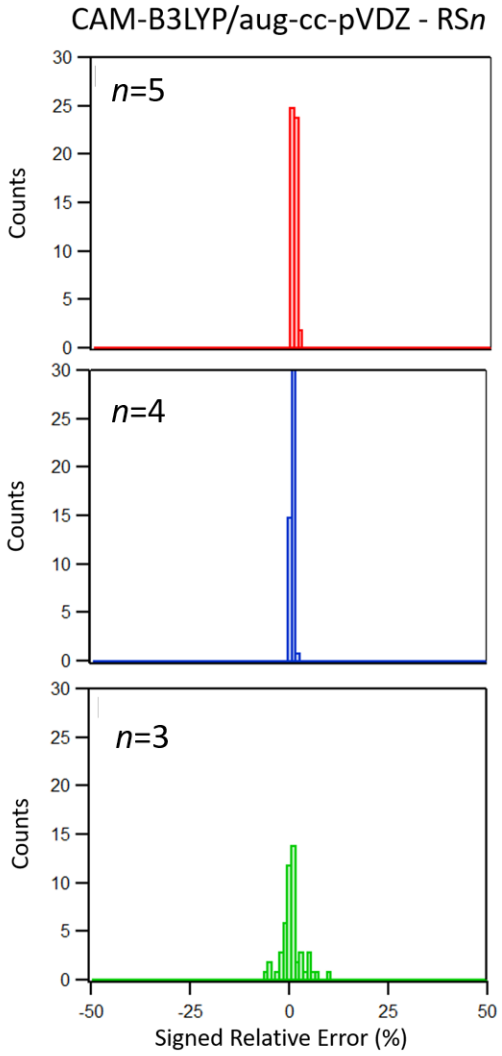


Figure A.3: Histogram of the signed relative errors for each of the 51 molecules. These results are calculated with CAM-B3LYP, using relative thresholds determined from $\langle|\mu|\rangle\langle m\rangle$.

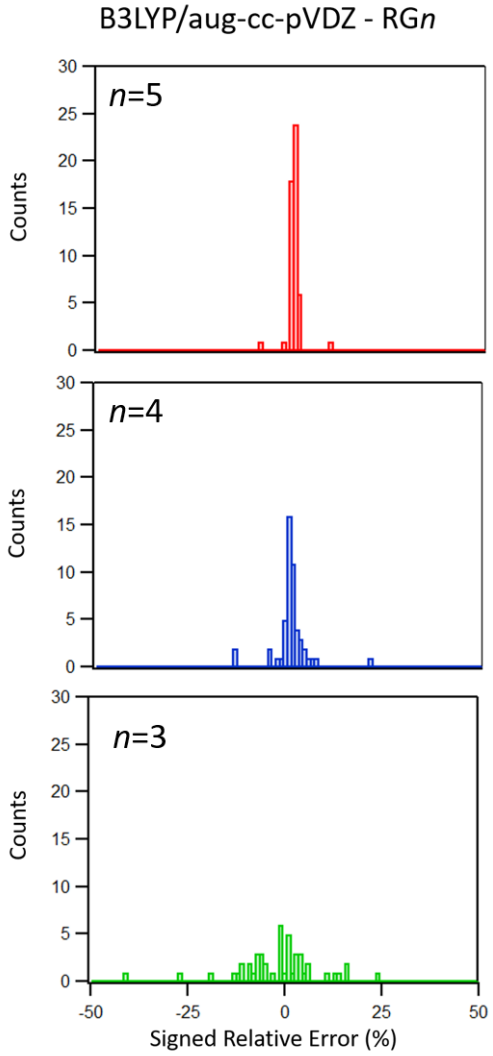


Figure A.4: Histogram of the signed relative errors for each of the 51 molecules. These results are calculated with B3LYP, using relative thresholds determined from $\langle m \rangle^2$.

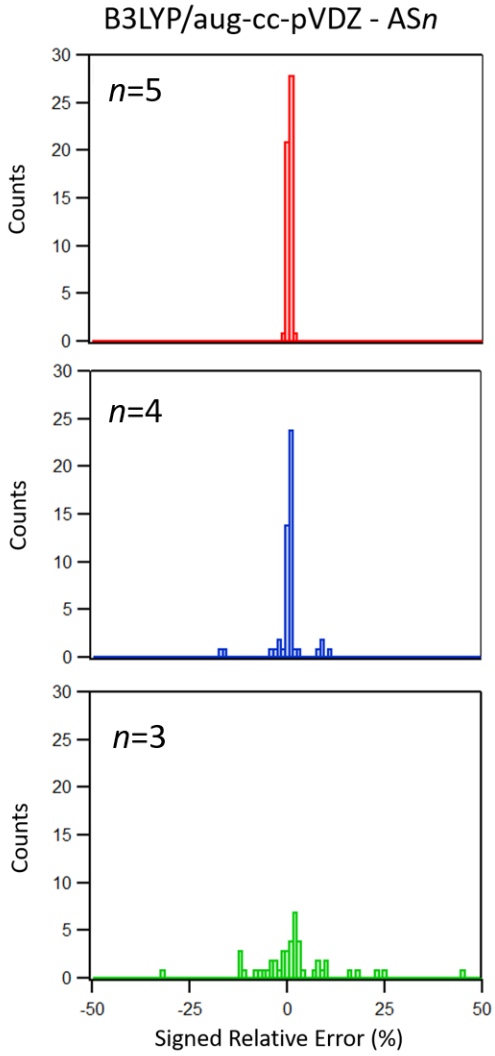


Figure A.5: Histogram of the signed relative errors for each of the 51 molecules. These results are calculated with B3LYP, using absolute thresholds determined from $\langle|\mu|\rangle\langle m\rangle$.

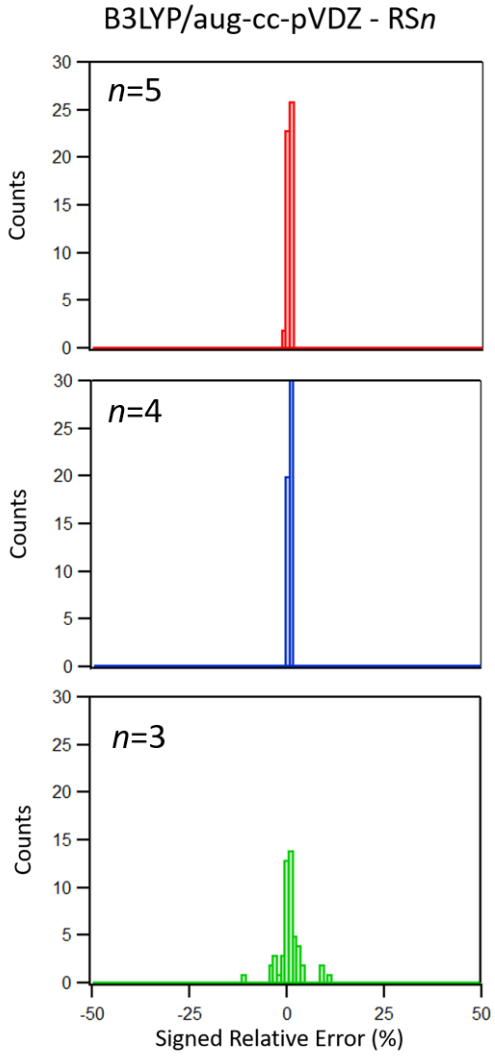


Figure A.6: Histogram of the signed relative errors for each of the 51 molecules. These results are calculated with B3LYP, using relative thresholds determined from $\langle|\mu|\rangle\langle m\rangle$.

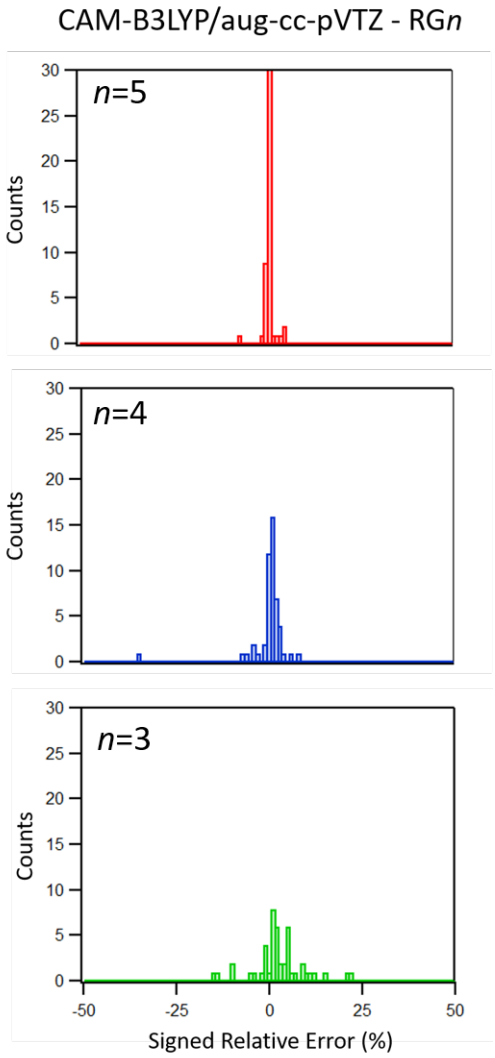


Figure A.7: Histogram of the signed relative errors for each of the 51 molecules. These results are calculated with CAM-B3LYP, using relative thresholds determined from $\langle m \rangle^2$.

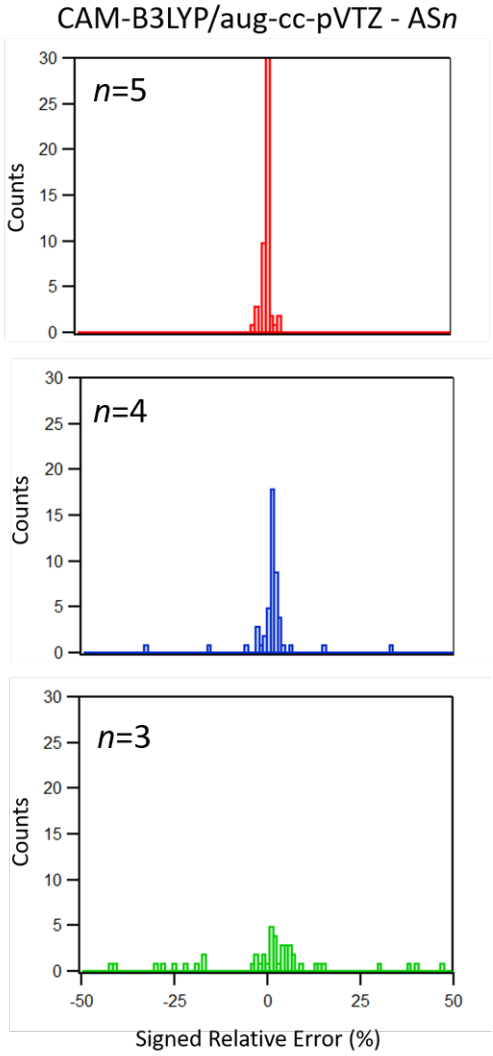


Figure A.8: Histogram of the signed relative errors for each of the 51 molecules. These results are calculated with CAM-B3LYP, using absolute thresholds determined from $\langle|\mu|\rangle\langle m\rangle$.

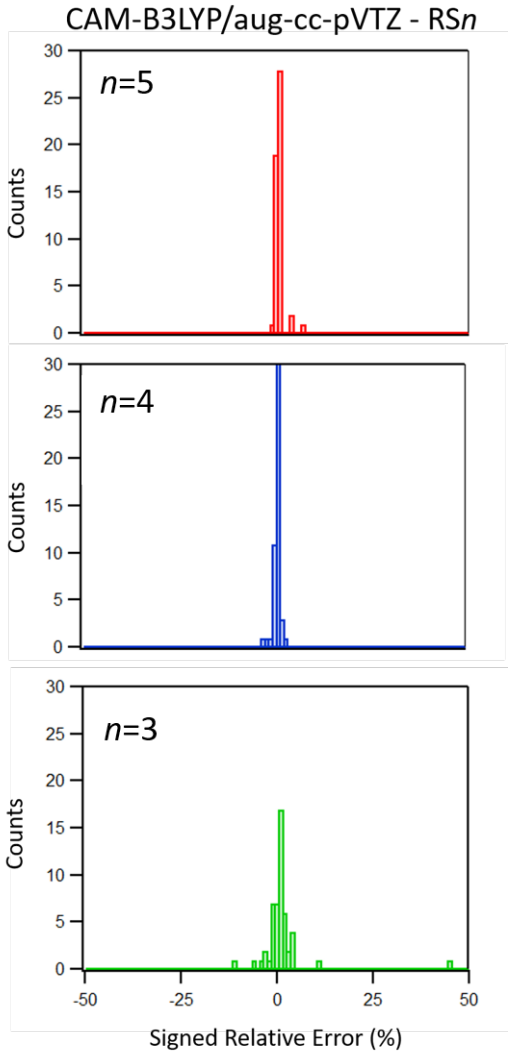


Figure A.9: Histogram of the signed relative errors for each of the 51 molecules. These results are calculated with CAM-B3LYP, using relative thresholds determined from $\langle|\mu|\rangle\langle m\rangle$.

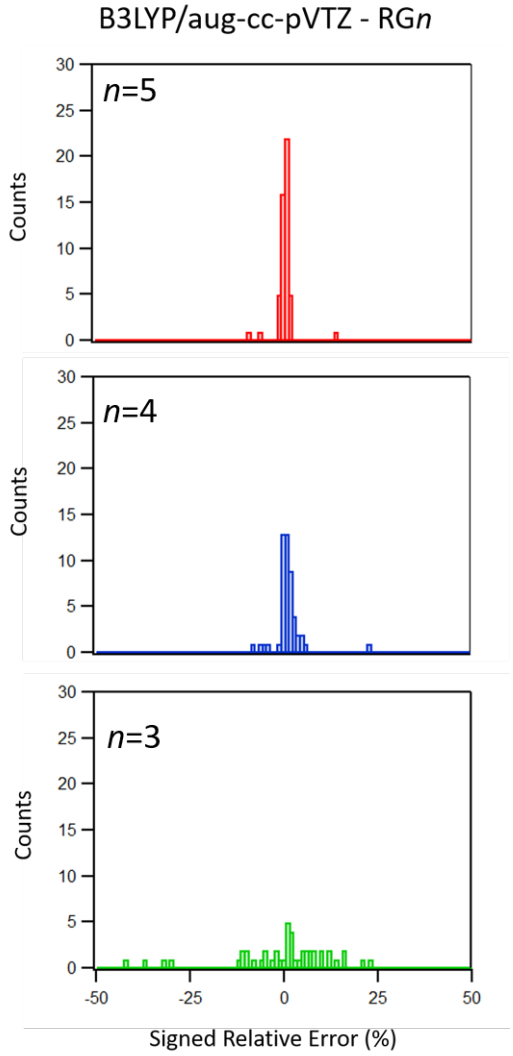


Figure A.10: Histogram of the signed relative errors for each of the 51 molecules. These results are calculated with B3LYP, using relative thresholds determined from $\langle m \rangle^2$.

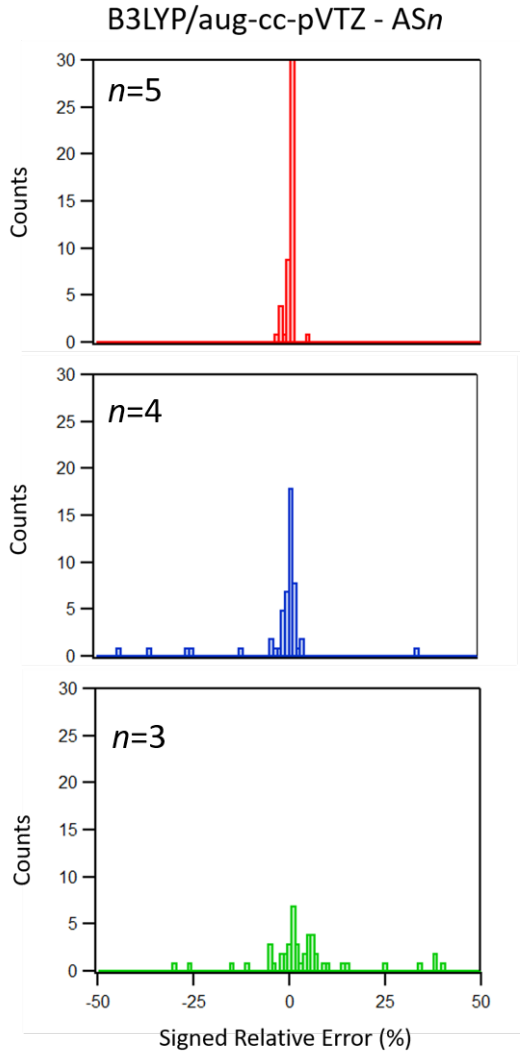


Figure A.11: Histogram of the signed relative errors for each of the 51 molecules. These results are calculated with B3LYP, using absolute thresholds determined from $\langle|\mu|\rangle\langle m\rangle$.

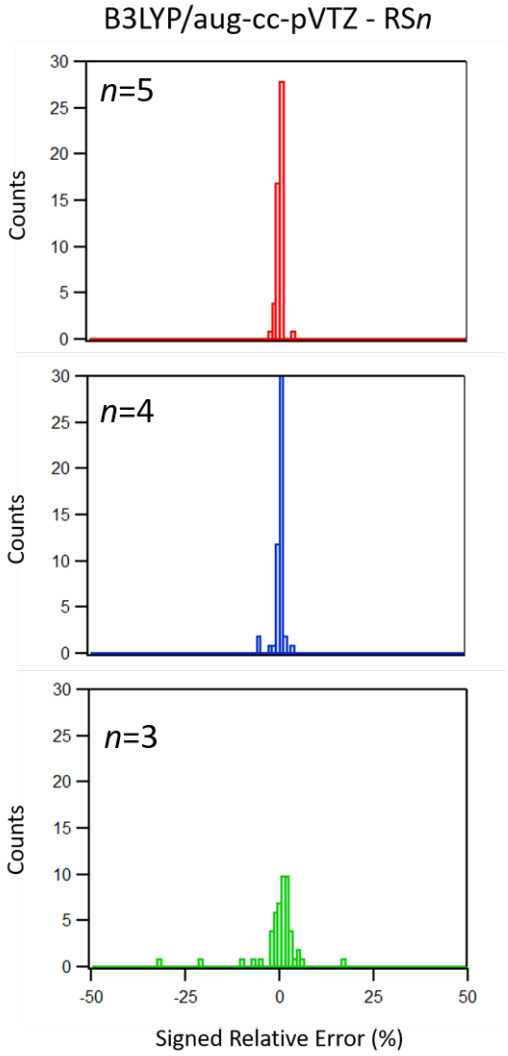


Figure A.12: Histogram of the signed relative errors for each of the 51 molecules. These results are calculated with B3LYP, using relative thresholds determined from $\langle|\mu|\rangle\langle m\rangle$.

Table A.17: Coordinates (Å) for molecule 43.

Element	X	Y	Z
C	0.576832	-0.529628	-0.988559
C	-0.803559	-0.994654	-0.742419
C	-1.741572	-0.284703	-0.107924
C	-1.447463	1.081562	0.458246
C	-0.235527	1.797133	-0.167669
C	0.899677	0.903896	-0.632089
C	1.554906	-0.205542	0.150626
H	1.039768	-0.941599	-1.877401
H	-1.056135	-1.983348	-1.111685
C	-3.128149	-0.812487	0.123927
H	-3.249446	-1.817423	-0.279462
H	-3.877272	-0.164560	-0.340904
H	-3.366975	-0.845140	1.191569
H	-2.331145	1.716434	0.344571
H	-1.307909	0.988011	1.540248
H	-0.582680	2.339015	-1.051417
H	0.134419	2.557951	0.524166
H	1.571126	1.412949	-1.314575
C	1.206709	-0.510337	1.591906
H	0.137399	-0.592670	1.761187
H	1.601895	0.260331	2.258922
H	1.655215	-1.460552	1.890866
C	3.021514	-0.439741	-0.158468
H	3.252928	-0.222437	-1.201919
H	3.301559	-1.477598	0.037994
H	3.657040	0.197639	0.462381

Table A.18: Coordinates (Å) for molecule 44.

Element	X	Y	Z
C	0.533673	-0.452885	-0.919589
C	-0.775127	-1.038823	-0.547619
C	-1.812783	-0.301049	-0.146244
C	-1.685170	1.202076	-0.020487
C	-0.266200	1.659183	0.355841
C	0.812750	0.960262	-0.460645
C	1.645692	-0.185246	0.084888
H	0.879519	-0.724264	-1.911920
H	-0.889367	-2.114513	-0.641961
C	-3.154739	-0.893706	0.167893
H	-3.168899	-1.971103	0.004808
H	-3.934392	-0.444709	-0.455131
H	-3.439458	-0.699741	1.206349
H	-1.985637	1.660610	-0.970160
H	-2.397628	1.569958	0.722358
H	-0.181909	2.738532	0.218547
H	-0.115598	1.480011	1.421702
H	1.326761	1.596386	-1.170612
C	1.500568	-0.633686	1.521830
H	0.463841	-0.660336	1.850012
H	2.053354	0.027547	2.194387
H	1.907427	-1.640234	1.644945
C	3.064166	-0.292128	-0.436791
H	3.134886	0.036322	-1.474607
H	3.746107	0.324621	0.154644
H	3.424017	-1.323081	-0.387817

Table A.19: Coordinates (Å) for molecule 45.

Element	X	Y	Z
C	-1.320087	-0.253322	0.015631
C	-0.030259	-1.067500	0.031545
C	1.077321	-0.083115	-0.371103
C	0.561094	1.268783	0.169294
C	-0.951884	1.227308	-0.067010
O	-2.439012	-0.698941	0.059449
H	0.119350	-1.413645	1.061010
H	-0.117619	-1.954404	-0.594544
H	1.091376	-0.020492	-1.464626
H	1.049911	2.118428	-0.306621
H	0.771512	1.332631	1.240763
H	-1.550913	1.814308	0.627569
H	-1.208378	1.566663	-1.075371
C	2.473062	-0.457556	0.109542
H	3.207043	0.294312	-0.186340
H	2.794582	-1.414944	-0.303568
H	2.499750	-0.538919	1.198745

Table A.20: Coordinates (Å) for molecule 46.

Element	X	Y	Z
C	1.209971	-0.168409	-0.037577
C	0.134587	-0.978918	0.680653
C	-1.173625	-0.194078	0.487730
C	-0.691218	1.277798	0.406650
C	0.650226	1.221105	-0.335091
O	2.314893	-0.557635	-0.321342
H	0.419876	-1.014664	1.737040
H	0.114995	-2.006666	0.320003
H	-1.834407	-0.324472	1.345474
H	-1.422475	1.922396	-0.080213
H	-0.538990	1.667400	1.415280
H	1.366302	1.989659	-0.047264
H	0.525901	1.298250	-1.418310
C	-1.926908	-0.630353	-0.771386
H	-2.827093	-0.030840	-0.915342
H	-2.228717	-1.676575	-0.703013
H	-1.312746	-0.526277	-1.668787

Table A.21: Coordinates (Å) for molecule 47.

Element	X	Y	Z
C	-1.085568	-0.389987	0.020590
C	0.443130	-0.402100	0.180898
C	-1.375441	1.113110	-0.032305
H	-1.540849	-0.853833	0.911689
H	-1.634867	1.498357	0.966937
H	-2.140984	1.390838	-0.764229
O	-0.146922	1.735140	-0.472878
N	0.861311	0.909354	0.106573
H	1.767953	1.125698	-0.294861
O	1.157082	-1.372740	0.350934
N	-1.385066	-1.075008	-1.238923
H	-0.930911	-1.986470	-1.252254
H	-2.385250	-1.232785	-1.333590

Table A.22: Coordinates (Å) for molecule 48.

Element	X	Y	Z
C	-1.444274	0.491851	-0.551958
C	0.077449	0.449126	-0.484982
C	0.643461	1.878362	-0.488549
C	-1.329102	2.711549	0.616347
C	-1.993653	1.330251	0.606122
H	1.730967	1.860265	-0.306701
H	0.369506	-0.008161	0.472052
H	-1.743691	0.948059	-1.508298
H	-1.654160	3.229037	-0.300682
H	-3.077083	1.476243	0.445413
O	0.114320	2.611887	0.595622
C	-1.688839	3.538932	1.838166
H	-1.211401	4.524491	1.776455
H	-2.777321	3.675327	1.891825
H	-1.352945	3.031037	2.748637
O	-1.769438	0.683559	1.855320
H	-2.017165	-0.243762	1.736355
O	-2.006171	-0.817637	-0.430640
H	-1.522549	-1.397423	-1.033769
O	0.517146	-0.348669	-1.585453
H	1.457266	-0.539852	-1.491487
O	0.355879	2.472018	-1.740669
H	0.747862	3.354163	-1.746906

Table A.23: Coordinates (Å) for molecule 49.

Element	X	Y	Z
C	-1.157368	-1.404668	0.304377
C	0.302282	-1.182628	-0.007987
C	0.873617	0.018913	-0.181836
C	0.076626	1.297772	-0.066214
C	-1.298352	1.094304	0.582529
C	-2.042723	-0.160570	0.068693
H	-1.537956	-2.255033	-0.282203
H	0.922528	-2.079140	-0.095203
H	0.655037	2.030616	0.519307
H	-1.158289	0.972892	1.668257
H	-0.027797	1.751084	-1.066741
H	-1.927194	1.982494	0.434258
H	-2.940501	-0.269816	0.693803
H	-1.256334	-1.713874	1.358909
C	2.329033	0.179208	-0.531685
H	2.854686	0.776700	0.230457
H	2.837489	-0.789270	-0.623882
H	2.441915	0.719871	-1.486030
C	-2.543384	-0.016108	-1.365960
C	-1.620514	-0.335622	-2.518733
H	-0.732952	0.310054	-2.525788
H	-1.247023	-1.367460	-2.454240
H	-2.141005	-0.218597	-3.477338
C	-3.803318	0.387966	-1.590478
H	-4.191627	0.520578	-2.601113
H	-4.489931	0.602273	-0.769775

Table A.24: Coordinates (Å) for molecule 50.

Element	X	Y	Z
C	0.586874	0.554700	0.323395
C	1.505734	1.596712	0.113545
C	0.992066	2.891629	0.009114
C	-0.387596	3.090099	0.101784
C	-1.208317	1.978161	0.305840
H	0.954830	-0.467242	0.428527
H	1.665507	3.737662	-0.140894
H	-0.820772	4.087359	0.023668
H	-2.291559	2.096042	0.385025
C	2.983688	1.312990	-0.028125
C	3.375392	0.672005	-1.383601
H	3.530297	2.277175	0.079899
C	4.802491	-0.050910	0.459171
C	4.598143	-0.227900	-1.055918
H	3.596288	1.444269	-2.130367
H	2.537877	0.075982	-1.764478
H	5.552225	0.746890	0.656417
H	5.145332	-0.961957	0.967572
H	5.494035	0.061961	-1.618413
H	4.380219	-1.275401	-1.296391
N	3.483310	0.342087	0.951646
C	3.484377	0.819335	2.323531
H	2.462553	1.068239	2.635451
H	3.863569	0.031599	2.987443
H	4.119994	1.721744	2.454766
N	-0.736998	0.727250	0.420204

Table A.25: Coordinates (Å) for molecule 51.

Element	X	Y	Z
C	-1.022567	0.558913	-0.212436
C	0.738164	2.201616	-0.519948
C	-0.607715	2.928478	-0.340386
C	-1.437069	1.886884	0.425571
H	-1.693635	0.262152	-1.029726
H	-1.004569	-0.249526	0.527004
H	1.295893	2.597092	-1.378378
H	-0.500944	3.880197	0.190547
H	-1.056102	3.116381	-1.324981
H	-2.520075	2.062230	0.343642
N	0.323993	0.824164	-0.742910
C	1.002394	-0.150739	-1.421835
C	2.358353	0.218885	-1.994116
H	3.024590	0.614148	-1.217325
H	2.795685	-0.682404	-2.431820
H	2.253365	0.981478	-2.780038
C	1.643010	2.426417	0.700442
O	0.510712	-1.267415	-1.568441
O	1.827684	3.516987	1.195323
O	2.256923	1.303535	1.127342
H	2.806788	1.550939	1.888541
O	-1.039136	1.825144	1.802659
H	-1.144470	2.697807	2.201773

Appendix B

Appendix for Small Basis Sets Optimized for Calculations of Optical Rotation

Table B.1: $[\alpha]_\omega$ [deg dm^{-1} (g/mL)$^{-1}$] values for the training set with aug-cc-pVDZ.

	CAM-B3LYP			B3LYP
Molecule	450 nm	589.3 nm[a]	633 nm	589.3 nm
1	-15397.83	-4926.46	-3894.96	-7000.30
2	13083.37	4567.17	3651.20	6364.12
3	-9968.41	-3629.52	-2917.46	-4822.08
4	-9809.52	-3540.40	-2843.39	-4822.80
5	-5925.32	-2387.07	-1948.04	-2913.12
6	-304.21	-160.21	-136.31	-169.40
7	-531.64	-212.58	-173.74	-254.35
8	792.76	388.21	326.10	407.89
9	-483.07	-226.91	-189.06	-347.72
10	-2213.73	-923.43	-757.55	-1192.48
11	300.21	169.19	145.50	165.25
12	-244.85	-136.99	-117.70	-134.75
13	235.26	128.21	109.75	149.84
14	218.58	126.14	108.85	131.22
15	127.53	90.25	79.65	136.37
16	285.30	156.37	134.02	175.88
17	-199.05	-104.57	-88.95	-96.04
18	-334.80	-176.19	-150.02	-214.64
19	-233.89	-118.95	-100.65	-145.99
20	117.61	70.86	61.68	76.00
21	156.22	86.58	74.32	102.12

a) The frequency and functional at which the optimization was performed.

Table B.2: $[\alpha]_\omega$ [deg dm^{-1} (g/mL)$^{-1}$] values for the training set with unoptimized augD-3-21G.

	CAM-B3LYP			B3LYP
Molecule	450 nm	589.3 nm[a]	633 nm	589.3 nm
1	-13323.89	-4462.56	-3548.37	-6260.98
2	11404.25	4135.84	3323.00	5713.16
3	-8791.68	-3305.89	-2669.01	-4361.77
4	-8679.77	-3237.31	-2611.47	-4381.92
5	-5298.30	-2186.56	-1790.50	-2654.70
6	-299.00	-158.28	-134.79	-165.30
7	-532.13	-216.57	-177.51	-255.91
8	767.74	379.03	318.86	394.87
9	-496.66	-233.71	-194.81	-353.13
10	-2158.14	-906.64	-744.61	-1163.98
11	266.29	148.63	127.61	146.74
12	-271.48	-150.98	-129.63	-148.64
13	239.06	130.60	111.85	151.89
14	188.06	108.89	94.01	113.25
15	157.57	106.46	93.51	155.71
16	292.78	160.44	137.50	180.32
17	-229.77	-122.04	-104.01	-114.70
18	-326.21	-171.25	-145.74	-210.78
19	-240.33	-122.18	-103.37	-149.95
20	126.56	75.33	65.44	79.34
21	157.86	87.01	74.63	102.86

a) The frequency and functional at which the optimization was performed.

Table B.3: $[\alpha]_\omega$ [deg dm^{-1} (g/mL)$^{-1}$] values for the training set with optimized augD-3-21G.

	CAM-B3LYP			B3LYP
Molecule	450 nm	589.3 nm[a]	633 nm	589.3 nm
1	-13690.04	-4567.78	-3630.52	-6394.13
2	11762.87	4252.78	3415.73	5856.18
3	-9024.49	-3386.24	-2733.20	-4453.59
4	-8922.50	-3321.46	-2678.78	-4476.48
5	-5468.00	-2254.57	-1846.07	-2724.90
6	-292.95	-154.84	-131.83	-163.34
7	-530.92	-211.99	-173.23	-252.04
8	787.26	389.23	327.53	403.04
9	-487.74	-229.11	-190.89	-352.66
10	-2213.59	-922.34	-756.60	-1189.05
11	294.54	164.90	141.70	156.98
12	-232.70	-129.56	-111.26	-126.93
13	241.87	131.95	112.98	153.74
14	228.81	132.11	114.05	130.00
15	123.19	92.60	82.28	139.83
16	295.95	162.20	139.01	182.95
17	-211.76	-111.96	-95.34	-103.96
18	-320.45	-169.26	-144.19	-207.14
19	-233.04	-117.77	-99.53	-146.70
20	124.09	73.76	64.07	75.37
21	160.55	88.96	76.36	103.67

a) The frequency and functional at which the optimization was performed.

Table B.4: $[\alpha]_\omega$ [deg dm^{-1} (g/mL)$^{-1}$] values for the control set with aug-cc-pVDZ.

	CAM-B3LYP			B3LYP
Molecule	450 nm	589.3 nm	633 nm	589.3 nm
1	-26.25	-18.65	-16.61	-17.13
2	-40.31	-24.52	-21.39	-13.36
3	-153.64	-86.89	-74.88	-83.11
4	-105.66	-48.53	-40.21	-65.78
5	-65.77	-38.86	-33.71	-41.32
6	4.30	2.32	1.98	1.89
7	31.33	9.95	7.64	8.46
8	107.35	42.24	34.19	58.18
9	-123.42	-57.31	-47.88	-67.11
10	154.47	85.82	73.71	94.08
11	180.56	95.97	81.80	116.21
12	139.24	80.36	69.46	84.02
13	126.84	68.21	58.25	79.69
14	75.28	44.72	38.83	43.04
15	50.03	28.48	24.56	29.63
16	174.61	94.34	80.63	111.35
17	118.41	71.76	62.28	54.61
18	109.26	58.93	50.35	45.71
19	31.03	6.68	4.29	26.44
20	-1.85	-9.52	-9.19	-8.97
21	167.92	97.83	84.01	90.67
22	98.13	57.04	49.34	58.50
23	28.59	15.31	13.02	11.77
24	-57.31	-38.42	-33.67	-24.97
25	-0.60	-23.53	-22.28	-10.03
26	137.75	67.88	57.30	80.78
27	-75.99	-39.09	-33.16	-39.09
28	402.50	201.15	169.66	201.15
29	535.24	235.89	195.39	235.89
30	-604.90	-274.59	-228.53	-274.59

Table B.5: $[\alpha]_\omega$ [deg dm^{-1} (g/mL)$^{-1}$] values for the control set with unoptimized augD-3-21G.

	CAM-B3LYP			B3LYP
Molecule	450 nm	589.3 nm	633 nm	589.3 nm
1	-27.79	-19.27	-17.11	-17.01
2	-70.02	-40.26	-34.80	-28.12
3	-186.04	-104.11	-89.56	-99.43
4	-112.06	-52.03	-43.22	-71.00
5	-55.46	-34.59	-30.24	-37.90
6	12.17	6.30	5.35	5.67
7	36.12	12.78	10.08	12.78
8	102.80	41.31	33.58	56.88
9	-131.84	-60.93	-50.86	-69.94
10	146.85	81.79	70.28	88.66
11	169.33	89.89	76.61	109.38
12	140.65	81.15	70.15	83.39
13	121.06	65.19	55.68	76.71
14	81.09	48.01	41.66	46.04
15	46.93	26.81	23.13	28.09
16	177.93	96.43	82.47	113.01
17	91.28	54.52	47.20	39.72
18	110.01	60.01	51.37	46.11
19	36.63	9.49	6.65	29.09
20	2.08	-7.84	-7.81	-8.11
21	175.90	101.78	87.41	95.29
22	87.72	50.84	43.95	53.22
23	44.43	22.48	18.95	20.43
24	-59.27	-39.46	-34.58	-25.39
25	9.13	-20.98	-20.40	-5.76
26	145.30	71.17	60.00	84.31
27	-95.76	-49.40	-41.92	-64.34
28	412.21	207.45	175.18	241.97
29	517.17	229.62	190.41	275.06
30	-597.63	-273.43	-227.84	-327.97

Table B.6: $[\alpha]_\omega$ [deg dm^{-1} (g/mL)$^{-1}$] values for the control set with optimized augD-3-21G.

	CAM-B3LYP			B3LYP
Molecule	450 nm	589.3 nm	633 nm	589.3 nm
1	-29.20	-19.83	-17.56	-15.28
2	-62.35	-36.88	-32.02	-27.51
3	-175.70	-99.07	-85.32	-95.70
4	-85.94	-37.91	-31.15	-55.82
5	-61.43	-37.31	-32.48	-40.91
6	6.66	3.23	2.70	3.10
7	46.33	17.70	14.23	16.84
8	97.18	36.72	29.47	53.46
9	-130.31	-59.35	-49.43	-68.25
10	155.70	86.79	74.59	93.43
11	177.01	94.36	80.47	114.45
12	137.61	79.52	68.76	82.13
13	118.74	63.99	54.67	75.29
14	74.10	43.95	38.16	42.73
15	33.18	19.09	16.48	19.17
16	169.71	91.79	78.46	108.08
17	111.96	66.78	57.86	49.86
18	100.49	54.68	46.78	39.42
19	37.89	10.26	7.31	29.47
20	11.00	-4.43	-5.06	-2.88
21	157.63	93.03	80.03	85.74
22	70.04	41.60	36.11	41.33
23	25.05	11.43	9.41	11.33
24	-75.20	-46.54	-40.47	-37.06
25	-21.80	-33.37	-30.65	-23.83
26	137.82	69.19	58.53	79.71
27	-75.95	-38.64	-32.71	-52.62
28	428.26	215.98	182.46	252.15
29	557.76	247.59	205.37	294.78
30	-611.85	-276.29	-229.76	-329.60

Table B.7: $[\alpha]_\omega$ [deg dm^{-1} (g/mL)$^{-1}$] values for the training set with aug-cc-pVTZ.

	CAM-B3LYP			B3LYP
Molecule	450 nm	589.3 nm[a]	633 nm	589.3 nm
1	-15545.08	-4957.5	-3917.93	-7046.15
2	13199.46	4594.76	3671.92	6402.94
3	-10044.4	-3648.62	-2931.87	-4847.45
4	-9906.93	-3565.88	-2862.82	-4857.38
5	-5955.15	-2394.97	-1954.02	-2922.54
6	-292.79	-153.8	-130.79	-161.34
7	-526.4	-211.62	-173.09	-252.29
8	802.02	393.35	330.52	412.68
9	-478.22	-224.57	-187.09	-343.56
10	-2174.66	-909.64	-746.47	-1172.15
11	246.27	140.91	121.42	135.5
12	-247.83	-139.1	-119.57	-137.67
13	227.85	124.23	106.36	144.67
14	225.69	129.68	111.84	134.52
15	112.66	81.65	72.24	126.46
16	281.76	154.45	132.38	173.83
17	-202.45	-106.39	-90.5	-98.53
18	-326.1	-171.43	-145.93	-209.62
19	-233.18	-118.84	-100.59	-145.7
20	96.09	58.62	51.12	62.77
21	143.57	79.66	68.4	93.39

a) The frequency and functional at which the optimization was performed.

Table B.8: $[\alpha]_\omega$ [deg dm^{-1} (g/mL)$^{-1}$] values for the training set with unoptimized augT3-3-21G.

	CAM-B3LYP			B3LYP
Molecule	450 nm	589.3 nm[a]	633 nm	589.3 nm
1	-12763.05	-4309.89	-3430.44	-6073.00
2	10960.72	3999.27	3215.75	5556.37
3	-8433.65	-3187.03	-2574.64	-4228.27
4	-8308.82	-3118.69	-2517.87	-4242.60
5	-5093.55	-2109.96	-1728.68	-2575.65
6	-295.24	-156.60	-133.40	-163.34
7	-532.36	-215.71	-176.68	-254.40
8	756.54	373.14	313.81	388.85
9	-506.73	-238.10	-198.42	-359.93
10	-2170.46	-912.38	-749.44	-1171.67
11	325.71	181.33	155.66	179.28
12	-227.87	-126.15	-108.20	-122.76
13	229.80	125.24	107.21	145.04
14	177.73	103.95	89.88	107.69
15	87.29	69.82	62.39	118.13
16	291.09	159.41	136.61	179.76
17	-217.15	-114.50	-97.45	-105.28
18	-341.13	-179.51	-152.85	-218.46
19	-254.77	-130.10	-110.16	-158.96
20	146.02	85.73	74.31	93.40
21	196.43	108.59	93.19	127.55

a) The frequency and functional at which the optimization was performed.

Table B.9: $[\alpha]_\omega$ [deg dm^{-1} (g/mL)$^{-1}$] values for the training set with augT4-3-21G.

	CAM-B3LYP			B3LYP
Molecule	450 nm	589.3 nm[a]	633 nm	589.3 nm
1	-14108.33	-4633.37	-3674.96	-6525.46
2	12119.58	4320.41	3463.31	5984.28
3	-9254.07	-3432.05	-2765.60	-4536.42
4	-9115.29	-3352.74	-2699.45	-4543.95
5	-5557.93	-2272.18	-1858.19	-2762.01
6	-305.57	-161.51	-137.50	-169.25
7	-539.69	-219.65	-180.05	-258.93
8	828.48	409.53	344.61	429.44
9	-516.71	-242.77	-202.30	-366.97
10	-2168.14	-908.50	-745.81	-1167.16
11	292.11	162.06	139.05	159.61
12	-239.30	-133.37	-114.53	-130.34
13	235.26	128.59	110.14	148.49
14	218.30	125.30	108.04	128.93
15	121.63	86.09	76.00	135.22
16	295.28	162.01	138.88	181.99
17	-208.38	-110.40	-94.05	-101.48
18	-339.43	-178.07	-151.54	-218.83
19	-233.16	-118.91	-100.66	-146.32
20	93.70	56.91	49.59	60.91
21	157.73	87.00	74.62	103.39

a) The frequency and functional at which the optimization was performed.

Table B.10: $[\alpha]_\omega$ [deg dm^{-1} (g/mL)$^{-1}$] values for the training set with optimized augT3-3-21G.

	CAM-B3LYP			B3LYP
Molecule	450 nm	589.3 nm[a]	633 nm	589.3 nm
1	-13573.29	-4519.81	-3591.38	-6337.86
2	11659.48	4213.00	3383.47	5814.73
3	-8949.43	-3354.84	-2707.41	-4426.62
4	-8828.61	-3283.62	-2647.83	-4440.05
5	-5359.12	-2202.87	-1802.61	-2674.95
6	-284.33	-149.66	-127.32	-157.49
7	-537.21	-214.84	-175.55	-255.19
8	798.72	395.34	332.73	412.77
9	-483.82	-227.62	-189.71	-347.15
10	-2239.14	-935.37	-767.59	-1201.97
11	234.28	130.53	112.04	124.38
12	-243.65	-135.12	-115.95	-134.35
13	238.68	130.53	111.82	151.84
14	213.31	122.26	105.39	128.13
15	108.98	82.80	73.58	131.20
16	287.72	157.62	135.08	177.31
17	-211.18	-111.29	-94.71	-102.55
18	-334.38	-176.45	-150.30	-214.23
19	-247.12	-125.70	-106.37	-153.39
20	98.42	59.63	51.94	62.91
21	143.37	79.35	68.11	95.19

a) The frequency and functional at which the optimization was performed.

Table B.11: $[\alpha]_\omega$ [deg dm^{-1} (g/mL)$^{-1}$] values for the control set with aug-cc-pVTZ.

	CAM-B3LYP			B3LYP
Molecule	450 nm	589.3 nm	633 nm	589.3 nm
1	-13.59	-11.28	-10.23	-9.61
2	-35.76	-21.97	-19.19	-11.34
3	-153.51	-86.86	-74.86	-84.00
4	-101.09	-46.76	-38.81	-62.12
5	-63.01	-36.84	-31.90	-39.23
6	3.36	1.87	1.61	1.53
7	29.54	9.49	7.31	8.77
8	105.61	42.32	34.38	58.11
9	-118.70	-54.46	-45.39	-63.73
10	148.82	82.67	71.00	90.70
11	175.15	93.10	79.36	112.60
12	138.01	79.58	68.79	83.63
13	122.38	65.79	56.18	76.72
14	74.25	44.03	38.22	42.95
15	48.72	27.69	23.87	28.94
16	168.46	91.05	77.83	106.94
17	70.93	46.51	40.74	28.41
18	108.92	58.59	50.03	45.22
19	26.62	4.34	2.29	23.47
20	-3.24	-9.94	-9.52	-9.38
21	169.96	98.51	84.55	91.38
22	89.47	52.20	45.17	52.73
23	2.21	1.31	1.09	-3.73
24	-64.47	-42.75	-37.43	-30.49
25	-2.21	-24.75	-23.36	-12.22
26	133.39	65.50	55.26	78.50
27	-77.95	-40.25	-34.17	-53.73
28	403.74	201.86	170.28	238.65
29	531.89	235.77	195.46	280.53
30	-597.44	-271.79	-226.26	-324.86

Table B.12: $[\alpha]_\omega$ [deg dm^{-1} (g/mL)$^{-1}$] values for the control set with unoptimized augT3-3-21G.

	CAM-B3LYP			B3LYP
Molecule	450 nm	589.3 nm	633 nm	589.3 nm
1	-43.53	-27.55	-24.15	-27.47
2	-77.27	-43.72	-37.68	-32.35
3	-208.01	-116.12	-99.84	-113.01
4	-104.82	-48.64	-40.40	-65.58
5	-51.51	-32.83	-28.79	-37.89
6	12.09	5.65	4.70	5.26
7	36.69	12.07	9.35	11.08
8	109.52	43.25	35.03	58.03
9	-124.77	-57.86	-48.33	-68.12
10	157.03	87.40	75.09	94.52
11	179.44	95.67	81.60	116.73
12	143.32	82.51	71.30	85.59
13	125.63	67.71	57.85	79.61
14	83.79	49.26	42.70	47.63
15	52.24	29.74	25.64	30.79
16	175.74	95.11	81.31	110.94
17	114.46	68.05	58.92	50.38
18	108.97	59.58	51.02	45.04
19	37.75	9.84	6.91	30.00
20	-0.65	-9.35	-9.10	-8.14
21	162.60	96.11	82.75	88.59
22	84.80	49.40	42.74	49.81
23	67.93	37.21	31.86	34.12
24	-11.33	-14.11	-12.93	1.31
25	63.15	3.30	-0.03	19.47
26	178.02	86.60	72.94	98.92
27	-85.05	-42.66	-36.01	-58.73
28	420.81	212.02	179.07	247.06
29	524.53	232.55	192.80	278.27
30	-582.32	-265.44	-221.05	-317.79

Table B.13: $[\alpha]_\omega$ [deg dm^{-1} (g/mL)$^{-1}$] values for the control set with augT4-3-21G.

	CAM-B3LYP			B3LYP
Molecule	450 nm	589.3 nm	633 nm	589.3 nm
1	-12.84	-10.36	-9.36	-8.81
2	-60.06	-35.21	-30.53	-22.35
3	-176.09	-99.04	-85.27	-95.03
4	-96.03	-44.03	-36.48	-60.78
5	-38.20	-24.02	-21.01	-28.17
6	4.45	1.79	1.45	1.33
7	51.55	19.63	15.76	19.10
8	115.19	47.49	38.81	63.66
9	-119.27	-55.06	-45.94	-64.26
10	154.85	86.30	74.17	93.84
11	173.24	92.23	78.64	112.15
12	142.97	82.46	71.28	85.58
13	118.09	63.52	54.25	74.32
14	76.73	45.40	39.39	43.41
15	51.64	29.51	25.47	30.43
16	171.67	93.15	79.68	108.24
17	97.54	58.21	50.42	40.41
18	108.73	58.81	50.27	45.55
19	31.93	6.84	4.37	26.22
20	3.60	-6.98	-7.07	-5.86
21	165.25	95.99	82.46	88.09
22	73.80	43.15	37.35	43.82
23	53.37	28.27	24.04	25.22
24	-36.86	-27.31	-24.18	-12.48
25	38.51	-6.92	-8.54	8.72
26	169.82	84.09	71.01	96.71
27	-93.12	-48.54	-41.28	-62.85
28	406.66	203.77	171.95	239.96
29	534.47	237.03	196.52	281.15
30	-597.74	-271.65	-226.10	-323.75

Table B.14: $[\alpha]_\omega$ [deg dm^{-1} (g/mL)$^{-1}$] values for the control set with optimized augT3-3-21G.

	CAM-B3LYP			B3LYP
Molecule	450 nm	589.3 nm	633 nm	589.3 nm
1	-13.83	-11.33	-10.27	-8.65
2	-61.36	-36.36	-31.58	-25.40
3	-180.80	-102.00	-87.86	-99.01
4	-104.46	-47.91	-39.69	-66.54
5	-69.35	-41.40	-35.96	-45.79
6	10.72	5.84	5.00	4.62
7	32.43	11.34	8.93	10.86
8	106.36	42.47	34.49	60.03
9	-128.10	-57.48	-47.74	-64.11
10	168.30	93.88	80.69	101.80
11	178.02	94.99	81.02	115.80
12	142.32	82.11	70.98	85.54
13	128.59	69.53	59.44	82.58
14	85.23	50.25	43.59	49.71
15	52.91	30.26	26.12	30.79
16	180.17	97.77	83.63	114.94
17	54.44	34.99	30.57	14.21
18	114.18	62.31	53.34	48.15
19	47.56	15.28	11.58	36.97
20	15.20	-1.98	-2.95	-1.57
21	193.49	112.30	96.52	106.14
22	88.09	51.30	44.39	53.28
23	28.34	13.00	10.73	9.99
24	-59.42	-37.77	-32.94	-25.57
25	2.39	-19.28	-18.53	-6.86
26	129.46	64.72	54.71	73.33
27	-80.53	-40.07	-33.78	-53.68
28	376.27	187.50	158.04	221.66
29	521.82	229.06	189.59	276.15
30	-595.38	-268.25	-222.96	-323.35

Table B.15: Exponents for augD-3-21G after optimization in GAUSSIAN ready form.

```
 -H
              S                   1                1.00
      1.82399900D-02    0.1000000000D+01
              P                   1                1.00
      1.01998400D-01    0.1000000000D+01
****
 -C
              S                   1                1.00
      1.76953300D-01    0.1000000000D+01
              P                   1                1.00
      1.92471800D-02    0.1000000000D+01
              D                   1                1.00
      1.98450700D-01    0.1000000000D+01
****

 -O
              S                   1                1.00
      7.70271300D-02    0.1000000000D+01
              P                   1                1.00
      4.94562400D-02    0.1000000000D+01
              D                   1                1.00
      1.98838400D-01    0.1000000000D+01
****
 -N
              S                   1                1.00
      3.17259000D-02    0.1000000000D+01
              P                   1                1.00
      9.60110400D-02    0.1000000000D+01
              D                   1                1.00
      4.99835900D-01    0.1000000000D+01
****
```

Table B.16: Exponents for augT3-3-21G after optimization in GAUSSIAN ready form.

```
 -H
        S                   1                1.00
   2.34107700D-02    0.1000000000D+01
        P                   1                1.00
   1.01673600D-01    0.1000000000D+01
****
 -C
        S                   1                1.00
   1.96481300D-01    0.1000000000D+01
        P                   1                1.00
   4.83775800D-02    0.1000000000D+01
        D                   1                1.00
   1.72094000D-01    0.1000000000D+01
****

 -O
        S                   1                1.00
   3.56880700D-02    0.1000000000D+01
        P                   1                1.00
   7.72737200D-02    0.1000000000D+01
        D                   1                1.00
   2.53354900D-01    0.1000000000D+01
****
 -N
        S                   1                1.00
   3.07858300D-02    0.1000000000D+01
        P                   1                1.00
   1.54403600D-01    0.1000000000D+01
        D                   1                1.00
   5.59682600D-01    0.1000000000D+01
****
```

Table B.17: Total number of basis functions for each molecule with augD-3-21G and aug-cc-pVDZ.

	Training Set			Control Set	
Molecule	augD-3-21G	aug-cc-pVDZ	Molecule	augD-3-21G	aug-cc-pVDZ
1	678	852	1	112	146
2	628	788	2	199	260
3	590	742	3	230	301
4	617	769	4	143	187
5	502	632	5	173	219
6	143	187	6	100	118
7	293	379	7	212	274
8	236	310	8	305	397
9	469	599	9	305	397
10	200	256	10	212	274
11	147	191	11	243	315
12	135	173	12	243	315
13	200	256	13	274	356
14	143	187	14	274	356
15	169	215	15	274	356
16	281	361	16	231	297
17	286	374	17	116	150
18	312	402	18	286	374
19	294	375	19	286	374
20	143	187	20	274	356
21	149	196	21	536	694
			22	155	205
			23	174	228
			24	166	214
			25	176	223
			26	135	173
			27	286	374
			28	286	374
			29	193	251
			30	193	251

Table B.18: Total number of basis functions for each molecule with augT3-3-21G and aug-cc-pVTZ.

	Training Set			Control Set	
Molecule	augT3-3-21G	aug-cc-pVTZ	Molecule	augT3-3-21G	aug-cc-pVTZ
1	678	1794	1	112	322
2	628	1656	2	199	575
3	590	1564	3	230	667
4	617	1600	4	143	414
5	502	1334	5	173	464
6	143	414	6	100	224
7	293	828	7	212	598
8	236	690	8	305	874
9	469	1288	9	305	874
10	200	552	10	212	598
11	147	418	11	243	690
12	135	372	12	243	690
13	200	552	13	274	782
14	143	414	14	274	782
15	169	460	15	274	782
16	281	782	16	231	644
17	286	828	17	116	326
18	312	874	18	286	828
19	294	805	19	286	828
20	143	414	20	274	782
21	149	437	21	536	1518
			22	155	460
			23	174	506
			24	166	464
			25	176	473
			26	135	372
			27	286	828
			28	286	828
			29	193	552
			30	193	552

Table B.19: Coordinates (Å) for molecule **15** of the training set.

Element	X	Y	Z
C	-1.085568	-0.389987	0.020590
C	0.443130	-0.402100	0.180898
C	-1.375441	1.113110	-0.032305
H	-1.540849	-0.853833	0.911689
H	-1.634867	1.498357	0.966937
H	-2.140984	1.390838	-0.764229
O	-0.146922	1.735140	-0.472878
N	0.861311	0.909354	0.106573
H	1.767953	1.125698	-0.294861
O	1.157082	-1.372740	0.350934
N	-1.385066	-1.075008	-1.238923
H	-0.930911	-1.986470	-1.252254
H	-2.385250	-1.232785	-1.333590

Table B.20: Coordinates (Å) for molecule **16** of the training set.

Element	X	Y	Z
C	-1.444274	0.491851	-0.551958
C	0.077449	0.449126	-0.484982
C	0.643461	1.878362	-0.488549
C	-1.329102	2.711549	0.616347
C	-1.993653	1.330251	0.606122
H	1.730967	1.860265	-0.306701
H	0.369506	-0.008161	0.472052
H	-1.743691	0.948059	-1.508298
H	-1.654160	3.229037	-0.300682
H	-3.077083	1.476243	0.445413
O	0.114320	2.611887	0.595622
C	-1.688839	3.538932	1.838166
H	-1.211401	4.524491	1.776455
H	-2.777321	3.675327	1.891825
H	-1.352945	3.031037	2.748637
O	-1.769438	0.683559	1.855320
H	-2.017165	-0.243762	1.736355
O	-2.006171	-0.817637	-0.430640
H	-1.522549	-1.397423	-1.033769
O	0.517146	-0.348669	-1.585453
H	1.457266	-0.539852	-1.491487
O	0.355879	2.472018	-1.740669
H	0.747862	3.354163	-1.746906

Table B.21: Coordinates (Å) for molecule **17** of the training set.

Element	X	Y	Z
C	-1.157368	-1.404668	0.304377
C	0.302282	-1.182628	-0.007987
C	0.873617	0.018913	-0.181836
C	0.076626	1.297772	-0.066214
C	-1.298352	1.094304	0.582529
C	-2.042723	-0.160570	0.068693
H	-1.537956	-2.255033	-0.282203
H	0.922528	-2.079140	-0.095203
H	0.655037	2.030616	0.519307
H	-1.158289	0.972892	1.668257
H	-0.027797	1.751084	-1.066741
H	-1.927194	1.982494	0.434258
H	-2.940501	-0.269816	0.693803
H	-1.256334	-1.713874	1.358909
C	2.329033	0.179208	-0.531685
H	2.854686	0.776700	0.230457
H	2.837489	-0.789270	-0.623882
H	2.441915	0.719871	-1.486030
C	-2.543384	-0.016108	-1.365960
C	-1.620514	-0.335622	-2.518733
H	-0.732952	0.310054	-2.525788
H	-1.247023	-1.367460	-2.454240
H	-2.141005	-0.218597	-3.477338
C	-3.803318	0.387966	-1.590478
H	-4.191627	0.520578	-2.601113
H	-4.489931	0.602273	-0.769775

Table B.22: Coordinates (Å) for molecule **18** of the training set.

Element	X	Y	Z
C	0.586874	0.554700	0.323395
C	1.505734	1.596712	0.113545
C	0.992066	2.891629	0.009114
C	-0.387596	3.090099	0.101784
C	-1.208317	1.978161	0.305840
H	0.954830	-0.467242	0.428527
H	1.665507	3.737662	-0.140894
H	-0.820772	4.087359	0.023668
H	-2.291559	2.096042	0.385025
C	2.983688	1.312990	-0.028125
C	3.375392	0.672005	-1.383601
H	3.530297	2.277175	0.079899
C	4.802491	-0.050910	0.459171
C	4.598143	-0.227900	-1.055918
H	3.596288	1.444269	-2.130367
H	2.537877	0.075982	-1.764478
H	5.552225	0.746890	0.656417
H	5.145332	-0.961957	0.967572
H	5.494035	0.061961	-1.618413
H	4.380219	-1.275401	-1.296391
N	3.483310	0.342087	0.951646
C	3.484377	0.819335	2.323531
H	2.462553	1.068239	2.635451
H	3.863569	0.031599	2.987443
H	4.119994	1.721744	2.454766
N	-0.736998	0.727250	0.420204

Table B.23: Coordinates (Å) for molecule **19** of the training set.

Element	X	Y	Z
C	-1.022567	0.558913	-0.212436
C	0.738164	2.201616	-0.519948
C	-0.607715	2.928478	-0.340386
C	-1.437069	1.886884	0.425571
H	-1.693635	0.262152	-1.029726
H	-1.004569	-0.249526	0.527004
H	1.295893	2.597092	-1.378378
H	-0.500944	3.880197	0.190547
H	-1.056102	3.116381	-1.324981
H	-2.520075	2.062230	0.343642
N	0.323993	0.824164	-0.742910
C	1.002394	-0.150739	-1.421835
C	2.358353	0.218885	-1.994116
H	3.024590	0.614148	-1.217325
H	2.795685	-0.682404	-2.431820
H	2.253365	0.981478	-2.780038
C	1.643010	2.426417	0.700442
O	0.510712	-1.267415	-1.568441
O	1.827684	3.516987	1.195323
O	2.256923	1.303535	1.127342
H	2.806788	1.550939	1.888541
O	-1.039136	1.825144	1.802659
H	-1.144470	2.697807	2.201773

Table B.24: Coordinates (Å) for molecule **27** of the control set.

Element	X	Y	Z
C	0.576832	-0.529628	-0.988559
C	-0.803559	-0.994654	-0.742419
C	-1.741572	-0.284703	-0.107924
C	-1.447463	1.081562	0.458246
C	-0.235527	1.797133	-0.167669
C	0.899677	0.903896	-0.632089
C	1.554906	-0.205542	0.150626
H	1.039768	-0.941599	-1.877401
H	-1.056135	-1.983348	-1.111685
C	-3.128149	-0.812487	0.123927
H	-3.249446	-1.817423	-0.279462
H	-3.877272	-0.164560	-0.340904
H	-3.366975	-0.845140	1.191569
H	-2.331145	1.716434	0.344571
H	-1.307909	0.988011	1.540248
H	-0.582680	2.339015	-1.051417
H	0.134419	2.557951	0.524166
H	1.571126	1.412949	-1.314575
C	1.206709	-0.510337	1.591906
H	0.137399	-0.592670	1.761187
H	1.601895	0.260331	2.258922
H	1.655215	-1.460552	1.890866
C	3.021514	-0.439741	-0.158468
H	3.252928	-0.222437	-1.201919
H	3.301559	-1.477598	0.037994
H	3.657040	0.197639	0.462381

Table B.25: Coordinates (Å) for molecule **28** of the control set.

Element	X	Y	Z
C	0.533673	-0.452885	-0.919589
C	-0.775127	-1.038823	-0.547619
C	-1.812783	-0.301049	-0.146244
C	-1.685170	1.202076	-0.020487
C	-0.266200	1.659183	0.355841
C	0.812750	0.960262	-0.460645
C	1.645692	-0.185246	0.084888
H	0.879519	-0.724264	-1.911920
H	-0.889367	-2.114513	-0.641961
C	-3.154739	-0.893706	0.167893
H	-3.168899	-1.971103	0.004808
H	-3.934392	-0.444709	-0.455131
H	-3.439458	-0.699741	1.206349
H	-1.985637	1.660610	-0.970160
H	-2.397628	1.569958	0.722358
H	-0.181909	2.738532	0.218547
H	-0.115598	1.480011	1.421702
H	1.326761	1.596386	-1.170612
C	1.500568	-0.633686	1.521830
H	0.463841	-0.660336	1.850012
H	2.053354	0.027547	2.194387
H	1.907427	-1.640234	1.644945
C	3.064166	-0.292128	-0.436791
H	3.134886	0.036322	-1.474607
H	3.746107	0.324621	0.154644
H	3.424017	-1.323081	-0.387817

Table B.26: Coordinates (Å) for molecule **29** of the control set.

Element	X	Y	Z
C	-1.320087	-0.253322	0.015631
C	-0.030259	-1.067500	0.031545
C	1.077321	-0.083115	-0.371103
C	0.561094	1.268783	0.169294
C	-0.951884	1.227308	-0.067010
O	-2.439012	-0.698941	0.059449
H	0.119350	-1.413645	1.061010
H	-0.117619	-1.954404	-0.594544
H	1.091376	-0.020492	-1.464626
H	1.049911	2.118428	-0.306621
H	0.771512	1.332631	1.240763
H	-1.550913	1.814308	0.627569
H	-1.208378	1.566663	-1.075371
C	2.473062	-0.457556	0.109542
H	3.207043	0.294312	-0.186340
H	2.794582	-1.414944	-0.303568
H	2.499750	-0.538919	1.198745

Table B.27: Coordinates (Å) for molecule **30** of the control set.

Element	X	Y	Z
C	1.209971	-0.168409	-0.037577
C	0.134587	-0.978918	0.680653
C	-1.173625	-0.194078	0.487730
C	-0.691218	1.277798	0.406650
C	0.650226	1.221105	-0.335091
O	2.314893	-0.557635	-0.321342
H	0.419876	-1.014664	1.737040
H	0.114995	-2.006666	0.320003
H	-1.834407	-0.324472	1.345474
H	-1.422475	1.922396	-0.080213
H	-0.538990	1.667400	1.415280
H	1.366302	1.989659	-0.047264
H	0.525901	1.298250	-1.418310
C	-1.926908	-0.630353	-0.771386
H	-2.827093	-0.030840	-0.915342
H	-2.228717	-1.676575	-0.703013
H	-1.312746	-0.526277	-1.668787